

AD-A102 868

ROYAL AIRCRAFT ESTABLISHMENT FARNBOROUGH (ENGLAND) F/6 12/1
NECESSARY AND SUFFICIENT CONDITIONS IN NONLINEAR OPTIMIZATION. (U)
JAN 81 J J SKROBANSKI

UNCLASSIFIED

RAE-TN-AERO-1888

DRIC-BR-77783

NL

1-4-1
AD-A102 868

END
DATE
FILMED
9-81
DTIC

TECH. MEMO
AERO 1888

BR77783

TECH. MEMO
AERO 1888 ✓

18 DRIC

19 BR-77183

LEVEL II

1

AD A102868

ROYAL AIRCRAFT ESTABLISHMENT

14 RAE-TM-AERO-1888

6

NECESSARY AND SUFFICIENT CONDITIONS IN NONLINEAR OPTIMIZATION

by

10 J. J./Skrobanski

11 Jand... 81

12 53

DTIC
ELECTE
S AUG 14 1981 D
E

7 Technical memo

DTIC FILE COPY

UNLIMITED

310450

81 8 14 015

ROYAL AIRCRAFT ESTABLISHMENT

Technical Memorandum Aero 1888

Received for printing 16 January 1981

NECESSARY AND SUFFICIENT CONDITIONS IN NONLINEAR OPTIMIZATION

by

J. J. Skrobanski

SUMMARY

These notes are directed at the newcomer to nonlinear programming for whom a thorough understanding of Lagrange multipliers, the Kuhn-Tucker conditions and the duality theorem is essential. The notes attempt to explain these foundations of the theory and what motivates them. Special cases of one or two dimensions are considered and are extended by means of the notation of vector differentiation to the case of n variables. The reader is taken in stages from the problem of unconstrained minimization, through the equation constrained problem, to the general constrained problem. The important Jacobian assumption is also discussed.

Copyright
©
Controller HMSO London
1981

Accession For	
NTIS	<input checked="checked" type="checkbox"/>
DTIC	<input checked="checked" type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	<input type="checkbox"/>
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

LIST OF CONTENTS

	<u>Page</u>
1 INTRODUCTION	3
2 PRELIMINARY THEORY	3
2.1 Notation	4
2.2 Vector differentiation	4
2.3 Tangent spaces and contours	7
3 THE UNCONSTRAINED PROBLEM	8
3.1 The one-dimensional case	8
3.2 The n-dimensional case	11
3.3 The quadratic unconstrained problem	13
4 THE EQUATION CONSTRAINED PROBLEM	17
4.1 The two-dimensional problem	17
4.2 The n-dimensional problem	20
4.3 Second order conditions	22
4.4 Implications of the Jacobian assumption	26
5 THE KUHN-TUCKER CONDITIONS	30
5.1 A one-dimensional problem	30
5.2 A special n-dimensional problem	31
5.3 Active constraints and the Jacobian assumption	32
5.4 The general constrained problem	33
5.5 Second order conditions	36
6 THE DUALITY THEOREM	36
6.1 The dual function	36
6.2 Statement of the theorem	37
6.3 Proof of part (i)	37
6.4 Proof of part (ii)	39
Appendix A Standard results of vector differentiation	43
Appendix B Statements of theorems assumed in the text	44
List of symbols	46
References	47
Illustrations	Figures 1-8
Report documentation page	inside back cover

1 INTRODUCTION

These notes are directed at the newcomer to optimization and nonlinear programming. Such a reader is confronted with a bewildering maze of conflicting and, in the author's opinion, inadequate notation. This is perhaps to be expected in one of the newest and most rapidly developing branches of mathematics, but it is a pity because the foundations of the subject can be made to appear deep and subtle, when in reality they consist of simple results that are easy to derive.

Optimization is concerned with the problem of minimizing a function of several (and often many) real-valued variables. If the variables themselves are restricted to satisfy other functional relations, the problem is said to be *constrained*. It should also be noted that if we are able to minimize a function f then we can also maximize the function $-f$, and vice-versa.

Nonlinear programming consists largely of a collection of algorithms for use by a computer to solve optimization problems that involve nonlinear functions. These algorithms are always iterative and, for unconstrained problems, the iterations are designed to converge to points that satisfy various *necessary and sufficient conditions*. In addition, for constrained problems the techniques of *Lagrange multipliers* and the *duality theorem* are required to help ensure the iterations converge successfully. A thorough knowledge of these foundations of optimization theory is thus essential before algorithms to solve practical problems can be written, efficiently implemented, or their results meaningfully interpreted.

In these notes we try not only to explain the foundations of the subject but also to show what motivates them, in the hope that this will increase the beginner's insight into the theory. We proceed by considering the special cases of functions of one or two variables and use geometrical interpretation to aid our understanding. The results thus obtained are then extended, by means of the notation of vector differentiation, to the case of functions of n variables, where the reader no longer has a geometrical crutch to rely on. The results obtained for the n -dimensional case bear a striking similarity to those for the simple case. It is hoped that this similarity will help to further increase the reader's understanding.

Also, these notes are deliberately structured to take the reader in stages from the comparatively simple problem of unconstrained minimization, through the *equation constrained* problem (sometimes called the *equality constrained* problem) to the *general constrained* problem (*ie* minimization subject to both equation and inequality constraints). However, it is shown that by employing the concept of *active constraints* the general constrained problem is dealt with by considering it as an equation constrained problem.

The important *Jacobian assumption* is also explained and the consequences of *not* assuming it to hold are discussed.

2 PRELIMINARY THEORY

In this section we mention some less well-known notation. The notation is adhered to throughout the rest of these notes. The reader should beware since other authors may

use different notation or they may use the same notation to denote different or even contradictory statements.

2.1 Notation

We denote the column vector of n variables x_1, \dots, x_n by \underline{x} . In these notes, underlined lower case letters will always denote column vectors. It will usually be made clear in the text whether the vectors are constants, variables or vector functions.

Matrices will sometimes be denoted by capital letters.

By

$$A = (a_{ij}) \quad \text{or} \quad a_{ij} = (A)_{ij}$$

we shall mean that A is the matrix whose i,j -entry is a_{ij} .

Let $A = (a_{ij})$ and $B = (b_{ij})$ be two $m \times n$ matrices. We shall write $A \leq B$ if and only if $a_{ij} \leq b_{ij}$ for all ij .

Suppose f is a function of n variables. Instead of writing $f = f(x_1, \dots, x_n)$ we shall frequently write $f = f(\underline{x})$ and say that f is a function of \underline{x} . Suppose f_1, \dots, f_m are m functions of \underline{x} . We can write this as $\underline{f}(\underline{x})$.

2.2 Vector differentiation

By vector differentiation we mean the differentiation of a function with respect to a vector. Note that the function can itself be a vector.

Let f be a function of n variables \underline{x} . Then if the partial derivatives

$\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ all exist define

$$\frac{df}{d\underline{x}} \triangleq \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

where the symbol \triangleq means that the left hand side is *defined* by the right hand side.

Note that many writers use the symbols ∇f , $\nabla_{\underline{x}} f$ or $\text{grad } f$ to denote vector differentiation (see, for instance Luenberger¹, Dixon²). However, with their notation it is sometimes not immediately clear which vector the function f is being differentiated with respect to. Also, with our present notation, many of the familiar results of scalar differentiation need little modification when extended to the case of vector differentiation. Thus the present notation is a useful memory aid and also provides good insight into how results are extended to more than three dimensions.

A few writers use the symbol $\partial f / \partial \underline{x}$ to denote vector differentiation (see Intrilligator³, from whose notation the present one has been modified). When we come to extend the concept of partial derivative to the vector case we shall see that this notation too can be confusing and inadequate.

Let \underline{f} be an $m \times 1$ vector function of \underline{x} . If the partial derivatives

$$\frac{\partial \underline{f}}{\partial x_i} = \begin{pmatrix} \frac{\partial f_1}{\partial x_i} \\ \vdots \\ \frac{\partial f_m}{\partial x_i} \end{pmatrix} \quad i = 1, \dots, n$$

exist, then define

$$\frac{d\underline{f}}{d\underline{x}} \triangleq \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_1} \\ \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_n} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

$\frac{d\underline{f}}{d\underline{x}}$ is called the *Jacobian matrix* of \underline{f} .

The definition of the second derivative of f , where f is now a scalar function, logically follows.

We define

$$\frac{d^2 f}{d\underline{x}^2} \triangleq \frac{d}{d\underline{x}} \left(\frac{d\underline{f}}{d\underline{x}} \right) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

$\frac{d^2 f}{d\underline{x}^2}$ is called the *Hessian matrix* of f .

With this notation, Taylor's series for functions of n variables \underline{x} is written

$$f(\underline{x} + \Delta \underline{x}) = f(\underline{x}) + \Delta \underline{x}^T \frac{df}{d\underline{x}} + \frac{1}{2!} \Delta \underline{x}^T \frac{d^2 f}{d\underline{x}^2} \Delta \underline{x} + \dots$$

It is straightforward to give $df/d\underline{x}$ a geometrical interpretation. The gradient of f at \underline{x}_0 along a direction \underline{v} is defined as

$$\lim_{h \rightarrow 0} \frac{f(\underline{x}_0 + h\underline{v}) - f(\underline{x}_0)}{h}.$$

Now from Taylor's series,

$$\frac{f(\underline{x}_0 + h\underline{v}) - f(\underline{x}_0)}{h} = \underline{v}^T \frac{df}{d\underline{x}} + \frac{1}{2!} h \underline{v}^T \frac{d^2 f}{d\underline{x}^2} \underline{v} + O(h^2) ,$$

so that the limit is $\underline{v}^T \frac{df}{d\underline{x}}$, which is the component of $df/d\underline{x}$ along \underline{v} . From elementary linear algebra we know that this is greatest when \underline{v} lies along $df/d\underline{x}$. Hence $df/d\underline{x}$ is the gradient of f along the line of steepest slope.

We now extend our notation to the case of *partial* differentiation with respect to vectors. From the theory of scalar partial differentiation, if f is a function of x_1, \dots, x_n then

$$df = dx_1 \frac{\partial f}{\partial x_1} + \dots + dx_n \frac{\partial f}{\partial x_n} = d\underline{x}^T \frac{df}{d\underline{x}} .$$

Suppose now that f is a function of two vectors \underline{x} and \underline{y} where \underline{y} is $m \times 1$. Then

$$\begin{aligned} df &= dx_1 \frac{\partial f}{\partial x_1} + \dots + dx_n \frac{\partial f}{\partial x_n} + dy_1 \frac{\partial f}{\partial y_1} + \dots + dy_m \frac{\partial f}{\partial y_m} \\ &= d\underline{x}^T \frac{\partial f}{\partial \underline{x}} + d\underline{y}^T \frac{\partial f}{\partial \underline{y}} , \end{aligned}$$

where we use curly ∂ to emphasise that differentiation is taking place with respect to only one of the possible *vector* variables.

The concept of total derivative can also be extended. Suppose the vector \underline{y} is a function of \underline{x} . If we keep all the independent variables except x_i , say, fixed and allow x_i to vary, then the dependent variables \underline{y} will also change. The total rate of change of f will then be given by

$$\begin{aligned} \left(\frac{\partial f}{\partial x_i} \right)_{\underline{x}} &= \left(\frac{\partial x_i}{\partial x_i} \right)_{\underline{x}} \left(\frac{\partial f}{\partial x_i} \right)_{\underline{x}, \underline{y}} + \left(\frac{\partial y_1}{\partial x_i} \right)_{\underline{x}} \frac{\partial f}{\partial y_1} + \dots + \left(\frac{\partial y_m}{\partial x_i} \right)_{\underline{x}} \frac{\partial f}{\partial y_m} \\ &= \left(\frac{\partial f}{\partial x_i} \right)_{\underline{x}, \underline{y}} + \left(\frac{\partial \underline{y}}{\partial x_i} \right)^T \frac{\partial f}{\partial \underline{y}} \end{aligned} \quad (2-2-1)$$

where the vector suffixes attached to the derivatives are a reminder that the derivatives with respect to x_i are not equal - the \underline{x} indicating, where it is present, that all the \underline{x} (except x_i) are kept fixed, the \underline{y} indicating that all the dependent variables are kept fixed. We then define the total derivative of f with respect to \underline{x} by

$$\frac{df}{d\underline{x}} \triangleq \frac{\partial f}{\partial \underline{x}} + \frac{d\underline{y}}{d\underline{x}} \frac{\partial f}{\partial \underline{y}} . \quad (2-2-2)$$

Equation (2-2-2) is of course obtained by repeating (2-2-1) for $i = 1, \dots, n$ and writing the result in vector form.

Many of the familiar standard results of scalar differentiation can be extended in a modified form to the vector case. Some of these results are used in subsequent sections. They are stated in Appendix A for the reader's convenience.

2.3 Tangent spaces and contours

When considering functions of two or three variables we can use our geometrical intuition to give us insight into the mathematical problem. This is reflected in the terminology we use. We say that

$$f(x, y, z) = 0 \quad (2-3-1)$$

represents a surface and that if the partial derivatives $\partial f/\partial x$, $\partial f/\partial y$ and $\partial f/\partial z$ are continuous, then the surface (2-3-1) is smooth. The vector

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)^T$$

is called the normal to the surface, and since equations of the form

$$ax + by + cz = \text{const}$$

represent planes, the equation

$$x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} + z \frac{\partial f}{\partial z} = x_0 \frac{\partial f}{\partial x} + y_0 \frac{\partial f}{\partial y} + z_0 \frac{\partial f}{\partial z}$$

must represent the tangent plane to the surface (2-3-1) at the point $(x_0, y_0, z_0)^T$.

Provided that the surface (2-3-1) is nowhere perpendicular to the (x, y) plane (i.e. $\partial f/\partial z$ is nowhere zero) our geometrical intuition tells us that we can draw contours of (2-3-1) onto the (x, y) plane of the form

$$g(x, y) = \text{const} . \quad (2-3-2)$$

Algebraically we do this first by transforming (2-3-1) into

$$h(x, y) = z$$

(our geometrical intuition suggests where this might *not* be possible) and then substituting constant values of z to obtain a family of contours like (2-3-2).

When our problem functions are of more than three variables, we no longer have a geometrical crutch to lean on, but the symbols we use look similar and so we employ a similar language. We say that

$$f(\underline{x}) = 0 \quad (2-3-3)$$

defines a surface whose tangent hyperplane or tangent space at \underline{x}_0 is given by

$$\underline{x}^T \frac{df}{d\underline{x}} = \underline{x}_0^T \frac{df}{d\underline{x}}$$

where $df/d\underline{x}$ is the normal to (2-3-3).

From the implicit function theorem (see Appendix B), provided $\partial f/\partial x_n$ is nowhere zero, we can rewrite (2-3-3) as

$$g(x_1, \dots, x_{n-1}) = x_n \quad (2-3-4)$$

which we interpret as a family of contours like

$$g(x_1, \dots, x_{n-1}) = \text{const}$$

setting $x_n = 0$ we obtain

$$g(x_1, \dots, x_{n-1}) = 0$$

which only underlines the obvious fact that the terms 'surface' and 'contour' are interchangeable. We shall use the term 'contour' in these notes.

Finally, we shall find it convenient to define a *path* from some starting point \underline{x}_0 to some endpoint \underline{x} , say, to be a sequence of points $\underline{x}_0, \underline{x}_1, \underline{x}_2, \dots$ which converge to \underline{x} .

3 THE UNCONSTRAINED PROBLEM

In this section we consider the unconstrained minimization problem

$U \quad \underset{\underline{x}}{\text{minimize}} \quad f(\underline{x})$
--

and we wish to obtain the necessary and sufficient conditions that \underline{x}^* be a solution of U . We derive these for the one-dimensional case first, in the hope that this will provide insight when we come to discuss the n -dimensional problem.

The only assumption we make is that $f(\underline{x})$ is continuously twice differentiable. This by no means restricts the scope of our theory since all practical problem functions can be approximated by polynomials that satisfy our assumption. We also restrict our definition of a minimum of f to *exclude* $-\infty$. This is not only convenient for us, but it also reflects the fact that iterative algorithms would fail to obtain such minima.

3.1 The one-dimensional case

We are interested in the one-dimensional problem

$U1 \quad \underset{x}{\text{minimize}} \quad f(x)$

By a solution of U1 we mean a real number x^* such that

$$f(x^* + \Delta x) \geq f(x^*) \quad (3-1-1)$$

for all small enough numbers Δx . Notice that this definition implies that $f(x^*)$ might only be a *local minimum* of f . In other words, there may be some other numbers \hat{x} satisfying $f(\hat{x}) < f(x^*)$ and our definition only ensures that they cannot be near to x^* . In particular there cannot be a path joining x^* to \hat{x} that a computer algorithm might follow and along which the value of f progressively decreases.

It is well known that the first order and second order necessary conditions for x^* to be a solution of U1 are, respectively

$$\frac{df}{dx}(x^*) = 0 \quad (3-1-2)$$

and

$$\frac{d^2f}{dx^2}(x^*) \geq 0 \quad (3-1-3)$$

whilst the sufficient conditions are

$$\frac{df}{dx}(x^*) = 0 \quad \text{and} \quad \frac{d^2f}{dx^2}(x^*) > 0. \quad (3-1-4)$$

These results are derived from Taylor's theorem

$$f(x^* + \Delta x) = f(x^*) + \Delta x \frac{df}{dx}(x^*) + \frac{1}{2!} \Delta x^2 \frac{d^2f}{dx^2}(x^*) + O(\Delta x^3). \quad (3-1-5)$$

Using (3-1-5) to eliminate $f(x^* + \Delta x)$ from (3-1-1) and taking $f(x^*)$ from each side gives

$$\Delta x \frac{df}{dx}(x^*) + \frac{1}{2} \Delta x^2 \frac{d^2f}{dx^2}(x^*) + O(\Delta x^3) \geq 0. \quad (3-1-6)$$

Suppose we set $\Delta x > 0$, then division of (3-1-6) by Δx gives

$$\frac{df}{dx}(x^*) + \frac{1}{2} \Delta x \frac{d^2f}{dx^2}(x^*) + O(\Delta x^2) \geq 0. \quad (3-1-7)$$

If we now let $\Delta x \rightarrow 0+$ we see that (3-1-7) implies

$$\frac{df}{dx}(x^*) \geq 0. \quad (3-1-8)$$

A similar process for $\Delta x < 0$ gives

$$\frac{df}{dx}(x^*) \leq 0. \quad (3-1-9)$$

(3-1-8) and (3-1-9) can only be both true if $\frac{df}{dx}(x^*) = 0$ and so we have the first result (3-1-2).

Now eliminating df/dx from (3-1-6) by using (3-1-2) leads to

$$\frac{1}{2}\Delta x^2 \frac{d^2f}{dx^2}(x^*) + O(\Delta x^3) \geq 0. \quad (3-1-10)$$

Dividing (3-1-10) by $\Delta x^2/2$ (which is always positive) leaves

$$\frac{d^2f}{dx^2}(x^*) + O(\Delta x) \geq 0. \quad (3-1-11)$$

On letting $\Delta x \rightarrow 0$ we see that $\frac{d^2f}{dx^2}(x^*) \geq 0$. This is the second order necessary condition (3-1-3).

To prove the sufficiency conditions (3-1-4) we assume that they hold and show that this implies $f(x^* + \Delta x) \geq f(x^*)$ for all sufficiently small Δx .

Now from the mean value theorem and the second mean value theorem, we know that there are numbers $\xi, \eta \in [0, 1]$ such that

$$\frac{df}{dx}(x^* + \xi\Delta x) = \frac{f(x^* + \Delta x) - f(x^*)}{\Delta x} \quad (3-1-12)$$

and

$$\frac{d^2f}{dx^2}(x^* + \eta\Delta x) = \frac{\frac{df}{dx}(x^* + \Delta x) - \frac{df}{dx}(x^*)}{\Delta x} \quad (3-1-13)$$

since, by hypothesis, $\frac{df}{dx}(x^*) = 0$, (3-1-13) implies that

$$\frac{d^2f}{dx^2}(x^* + \eta\Delta x) = \frac{1}{\Delta x} \frac{df}{dx}(x^* + \Delta x). \quad (3-1-14)$$

Now $\frac{d^2f}{dx^2}(x^*) > 0$ means that df/dx is strictly increasing, at least near to x^* .

Since $\frac{df}{dx}(x^*) = 0$ and $f(x)$ is continuously differentiable we must have from (3-1-14) that

$$\frac{df}{dx}(x^* + \xi\Delta x) > 0 \quad \text{for } \Delta x > 0 \quad (3-1-15)$$

and

$$\frac{df}{dx}(x^* + \xi \Delta x) < 0 \quad \text{for } \Delta x < 0 \quad (3-1-16)$$

for small enough Δx .

We can write (3-1-15) and (3-1-16) together as

$$\Delta x \frac{df}{dx}(x^* + \xi \Delta x) > 0 \quad (3-1-17)$$

but (3-1-17) is just the left hand side of (3-1-12) multiplied by Δx . Therefore, the right hand side of (3-1-12) multiplied by Δx is

$$f(x^* + \Delta x) - f(x^*) > 0$$

and so x^* is a solution of U .

3.2 The n-dimensional case

The necessary and sufficient conditions for the n-dimensional problem U can be derived in a similar manner. For the reader's convenience we state them first. We stress once again that they are only conditions for $f(\underline{x}^*)$ to be a *local* minimum of f . The first order and second order necessary conditions for \underline{x}^* to be a (local) solution of U are respectively

$$\frac{df}{d\underline{x}}(\underline{x}^*) = \underline{0} \quad (3-2-1)$$

and

$$\underline{\Delta x}^T \frac{d^2 f}{d\underline{x}^2} \underline{\Delta x} \geq 0 \quad (3-2-2)$$

for all small enough vectors $\underline{\Delta x}$. The sufficient conditions are

$$\frac{df}{d\underline{x}}(\underline{x}^*) = \underline{0} \quad \text{and} \quad \underline{\Delta x}^T \frac{d^2 f}{d\underline{x}^2} \underline{\Delta x} > 0. \quad (3-2-3)$$

The reader should be immediately aware of the similarity of these conditions with the one-dimensional case. They are also derived in a similar manner. The Taylor series in n dimensions gives

$$f(\underline{x}^* + \underline{\Delta x}) = f(\underline{x}^*) + \underline{\Delta x}^T \frac{df}{d\underline{x}}(\underline{x}^*) + \frac{1}{2!} \underline{\Delta x}^T \frac{d^2 f}{d\underline{x}^2}(\underline{x}^*) \underline{\Delta x} + \dots \quad (3-2-4)$$

Writing $\underline{\Delta x} = \Delta x \hat{u}$ where \hat{u} is the unit vector in the direction of $\underline{\Delta x}$ and Δx is the magnitude of $\underline{\Delta x}$, (3-2-4) becomes

$$f(\underline{x}^* + \underline{\Delta x}) = f(\underline{x}^*) + \Delta x \hat{u}^T \frac{df}{d\underline{x}}(\underline{x}^*) + \frac{1}{2!} \Delta x^2 \hat{u}^T \frac{d^2 f}{d\underline{x}^2}(\underline{x}^*) \hat{u} + \dots \quad (3-2-5)$$

To derive the necessary conditions we assume \underline{x}^* is a solution of U , i.e.

$$f(\underline{x}^* + \underline{\Delta x}) \geq f(\underline{x}^*) \quad (3-2-6)$$

for all vectors $\underline{\Delta x}$. Substituting this inequality into (3-2-5) gives

$$0 \leq \Delta x \hat{u}^T \frac{df}{d\mathbf{x}}(\mathbf{x}^*) + \frac{1}{2} \Delta x^2 \hat{u}^T \frac{d^2 f}{d\mathbf{x}^2}(\mathbf{x}^*) \hat{u} + O(\Delta x^3). \quad (3-2-7)$$

As in the one-dimensional case, we divide (3-2-7) by $\Delta x > 0$ and let $\Delta x \rightarrow 0+$ to obtain

$$\hat{u}^T \frac{df}{d\mathbf{x}}(\mathbf{x}^*) \leq 0. \quad (3-2-8)$$

If we now return to (3-2-4) and make the substitution $\Delta \mathbf{x} = \Delta x \hat{u}$ where \hat{u} is now the unit vector in the *opposite* direction to $\Delta \mathbf{x}$, so that the magnitude of $\Delta \mathbf{x}$ is $-\Delta x$, we can obtain, in the obvious way

$$\hat{u}^T \frac{df}{d\mathbf{x}}(\mathbf{x}^*) \geq 0. \quad (3-2-9)$$

(3-2-8) and (3-2-9) can only hold if

$$\hat{u}^T \frac{df}{d\mathbf{x}}(\mathbf{x}^*) = 0. \quad (3-2-10)$$

If we let \hat{u} run through the co-ordinate vectors \hat{e}_i in turn, we see that (3-2-10) implies $\frac{\partial f}{\partial x_i}(\mathbf{x}^*) = 0$ for all i , and hence $\frac{df}{d\mathbf{x}}(\mathbf{x}^*) = \mathbf{0}$. This is the first order necessary condition as required.

Because $\frac{df}{d\mathbf{x}}(\mathbf{x}^*) = \mathbf{0}$, (3-2-7) becomes

$$0 \leq \frac{1}{2} \Delta x^2 \hat{u}^T \frac{d^2 f}{d\mathbf{x}^2}(\mathbf{x}^*) \hat{u} + O(\Delta x^3). \quad (3-2-11)$$

Dividing by $\frac{1}{2} \Delta x^2$ and letting $\Delta x \rightarrow 0$ we see that

$$\hat{u}^T \frac{d^2 f}{d\mathbf{x}^2}(\mathbf{x}^*) \hat{u} \geq 0 \quad (3-2-12)$$

for all unit vectors \hat{u} and hence for all vectors $\Delta \mathbf{x} = \Delta x \hat{u}$. Thus we have proved the second order necessary condition.

All that now remains is to verify the sufficient conditions (3-2-3). From the mean value theorem in n dimensions and the second mean value theorem in n dimensions (see Appendix B) we know that there is a number $\xi \in [0, 1]$ such that

$$\Delta \mathbf{x}^T \frac{df}{d\mathbf{x}}(\mathbf{x}^* + \xi \Delta \mathbf{x}) = f(\mathbf{x}^* + \Delta \mathbf{x}) - f(\mathbf{x}^*) \quad (3-2-13)$$

and for all $n > 0$,

$$\frac{df}{d\mathbf{x}}(\mathbf{x}^* + \eta\Delta\mathbf{x}) = \frac{df}{d\mathbf{x}}(\mathbf{x}^*) + \eta \frac{d^2f}{d\mathbf{x}^2}(\mathbf{x}^*)\Delta\mathbf{x} + O(\eta^2) . \quad (3-2-14)$$

Noting that, by hypothesis, $\frac{df}{d\mathbf{x}}(\mathbf{x}^*) = \mathbf{0}$ and premultiplying (3-2-14) by $\Delta\mathbf{x}^T$ we have

$$\Delta\mathbf{x}^T \frac{df}{d\mathbf{x}}(\mathbf{x}^* + \eta\Delta\mathbf{x}) = \eta\Delta\mathbf{x}^T \frac{d^2f}{d\mathbf{x}^2}(\mathbf{x}^*)\Delta\mathbf{x} + O(\eta^2) . \quad (3-2-15)$$

The right hand side of (3-2-15) is greater than 0 (at least if η and $\Delta\mathbf{x}$ are small enough) since $\frac{d^2f}{d\mathbf{x}^2}(\mathbf{x}^*)$ is positive definite (i.e. $\Delta\mathbf{x}^T \frac{d^2f}{d\mathbf{x}^2}(\mathbf{x}^*)\Delta\mathbf{x} > 0$ for all $\Delta\mathbf{x} \neq \mathbf{0}$, again by hypothesis). Therefore the left hand side of (3-2-15) is also greater than 0, i.e.

$$\Delta\mathbf{x}^T \frac{df}{d\mathbf{x}}(\mathbf{x}^* + \eta\Delta\mathbf{x}) > 0 \quad (3-2-16)$$

for all (small enough) $\eta > 0$. By considering (3-2-13) we see that (3-2-16) implies $f(\mathbf{x}^* + \Delta\mathbf{x}) - f(\mathbf{x}^*) > 0$. We have thus shown that the conditions (3-2-3) are sufficient for \mathbf{x}^* to be a solution of U .

3.3 The quadratic unconstrained problem

We end this section by discussing the special case of U when f is a quadratic function of \mathbf{x} . The quadratic problem is important because many practical functions can be approximated by quadratic functions, at least close to their minimum \mathbf{x}^* . The well-known least squares method of solving

$$A\mathbf{x} = \mathbf{b}$$

by writing

$$\boldsymbol{\varepsilon} = A\mathbf{x} - \mathbf{b}$$

and minimising $\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}$ is an example of a quadratic problem. The quadratic problem also serves as a useful illustration of the general case.

The general quadratic problem is

$$\text{GQ} \quad \underset{\mathbf{x}}{\text{minimize}} \left(f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T A \mathbf{x} - \mathbf{b}^T \mathbf{x} + c \right)$$

There are several simplifying assumptions that we can make. First of all without loss of generality, A can be replaced by a symmetric matrix. Secondly, since

$$\min_{\mathbf{x}} f(\mathbf{x}) = \min_{\mathbf{x}} \{f(\mathbf{x}) - c\} + c$$

we can set $c = 0$. So we shall consider the problem

$$Q \underset{\underline{x}}{\text{minimize}} \left(f(\underline{x}) = \frac{1}{2} \underline{x}^T A \underline{x} - \underline{b}^T \underline{x} \right)$$

where A is symmetric.

We derive a theorem that tells us something of the conditions under which the problem Q has got a solution. The theorem's implications for the two-dimensional case are then more fully explored with the hope that the discussion will increase the reader's understanding of the theorem. However, we first introduce some preliminary definitions (see Kreyszig⁴).

Let A be an $n \times n$ matrix. A number λ which satisfies the equation

$$A \underline{x} = \lambda \underline{x} \quad (3-3-1)$$

for at least one non-zero vector \underline{x} is called an *eigenvalue* of A . The non-zero vectors \underline{x} which satisfy (3-3-1) are called the *eigenvectors* corresponding to λ . It can be shown that if A is a real symmetric matrix then all its eigenvalues $\lambda_1, \dots, \lambda_n$ are real (though not necessarily distinct) and that n corresponding *orthogonal* eigenvectors $\underline{x}_1, \dots, \underline{x}_n$ can be chosen.

The *rank* of a matrix is the number of its linearly independent columns. Let A be an $n \times n$ matrix of rank r . Then it can be shown that A has exactly $n - r$ zero eigenvalues.

An $n \times n$ matrix A is said to be *positive definite* if

$$\underline{x}^T A \underline{x} > 0 \quad (3-3-2)$$

for all non-zero vectors \underline{x} . If the strict inequality sign $>$ in (3-3-2) is replaced by \geq then A is said to be positive *semi-definite*. It can be proved that if A is a symmetric $n \times n$ positive semi-definite matrix of rank r then n eigenvalues $\lambda_1, \dots, \lambda_n$ exist, exactly r of which are positive (but not necessarily distinct) and the remaining $n - r$ eigenvalues are all zero. It should be clear from the above that, corresponding to the λ_i , n orthogonal eigenvectors $\underline{x}_1, \dots, \underline{x}_n$ can also be chosen. We are now in a position to state and prove our theorem.

Theorem Let A be a positive semi-definite matrix of rank r . Let $\underline{x}_1, \dots, \underline{x}_{n-r}$ be the orthogonal eigenvectors of A corresponding to zero eigenvalues. Then the problem Q has a solution if and only if the vector \underline{b} is orthogonal to every linear combination of $\underline{x}_1, \dots, \underline{x}_{n-r}$.

Proof Since A is positive semi-definite, there exists a vector $\underline{x}_0 \neq \underline{0}$ such that $\underline{x}_0^T A \underline{x}_0 = 0$.

Therefore

$$f(\underline{x}_0) = -\underline{b}^T \underline{x}_0 = \beta, \quad \text{say.}$$

By considering, if necessary, $-\underline{x}_0$ we can set $\beta \leq 0$, without loss of generality. But if $\beta < 0$ then $f(r\underline{x}_0) = r\beta + -\infty$ as $r \rightarrow \infty$. Hence Q has no solution. (Except for the excluded case $f(\underline{x}^*) = -\infty$.) Let $\underline{x}_{n-r+1}, \dots, \underline{x}_n$ be the remaining orthogonal eigenvectors of A with positive eigenvalues $\lambda_{n-r+1}, \dots, \lambda_n$. Now we may write $\underline{x}_0 = \alpha_1 \underline{x}_1 + \dots + \alpha_n \underline{x}_n$ for some unique $\alpha_1, \dots, \alpha_n$ since the eigenvectors form a basis.

Therefore

$$\begin{aligned}\underline{x}_0^T A \underline{x}_0 &= \underline{x}_0^T (\alpha_1 \lambda_1 \underline{x}_1 + \dots + \alpha_n \lambda_n \underline{x}_n) \\ &= \underline{x}_0^T (\alpha_{n-r+1} \lambda_{n-r+1} \underline{x}_{n-r+1} + \dots + \alpha_n \lambda_n \underline{x}_n)\end{aligned}$$

since $\lambda_1 = \dots = \lambda_{n-r} = 0$.

$$\begin{aligned}\text{Therefore } \underline{x}_0^T A \underline{x}_0 &= \left(\alpha_1 \underline{x}_1^T + \dots + \alpha_n \underline{x}_n^T \right) (\alpha_{n-r+1} \lambda_{n-r+1} \underline{x}_{n-r+1} + \dots + \alpha_n \lambda_n \underline{x}_n) \\ &= \alpha_{n-r+1}^2 \lambda_{n-r+1} \underline{x}_{n-r+1}^T \underline{x}_{n-r+1} + \dots + \alpha_n^2 \lambda_n \underline{x}_n^T \underline{x}_n\end{aligned}\quad (3-3-3)$$

since the \underline{x}_i are orthogonal. But the right hand side of (3-3-3) is positive if at least one $\alpha_i \neq 0$, ($i = n-r+1, \dots, n$). Hence $\alpha_i = 0$, ($i = n-r+1, \dots, n$) since $\underline{x}_0^T A \underline{x}_0 = 0$.

Therefore

$$\underline{x}_0 = \alpha_1 \underline{x}_1 + \dots + \alpha_{n-r} \underline{x}_{n-r+1}.$$

Thus \underline{x}_0 is a linear combination of $\underline{x}_1, \dots, \underline{x}_{n-r+1}$. Thus any vector \underline{x}_0 such that $\underline{x}_0^T A \underline{x}_0 = 0$ must be a linear combination of $\underline{x}_1, \dots, \underline{x}_{n-r+1}$. We have shown that if $\underline{b}^T \underline{x}_0 \neq 0$ where \underline{x}_0 is a linear combination of $\underline{x}_1, \dots, \underline{x}_{n-r}$ then the problem Q has no solution. We have thus proved the first part of the theorem: Q has a solution only if \underline{b} is orthogonal to every linear combination of $\underline{x}_1, \dots, \underline{x}_{n-r}$. Note that the case $\underline{b} = \underline{0}$ always has a solution $\underline{x}^* = \underline{0}$.

We next suppose that \underline{b} is orthogonal to every linear combination of $\underline{x}_1, \dots, \underline{x}_{n-r}$. We can write

$$\underline{x} = \alpha_1 \underline{x}_1 + \dots + \alpha_n \underline{x}_n$$

and

$$\underline{b} = \beta_1 \underline{x}_1 + \dots + \beta_n \underline{x}_n.$$

Using the above arguments we have

$$\underline{x}^T A \underline{x} = \alpha_{n-r+1}^2 \lambda_{n-r+1} \underline{x}_{n-r+1}^T \underline{x}_{n-r+1} + \dots + \alpha_n^2 \lambda_n \underline{x}_n^T \underline{x}_n.$$

Without loss of generality, we can assume the \underline{x}_i are orthonormal. We get

$$\underline{x}^T A \underline{x} = \alpha_{n-r+1}^2 \lambda_{n-r+1} + \dots + \alpha_n^2 \lambda_n.$$

But since \underline{b} is orthogonal to every linear combination of $\underline{x}_1, \dots, \underline{x}_{n-r}$,

$$\underline{b}^T \underline{x} = \alpha_{n-r+1} \beta_{n-r+1} + \dots + \alpha_n \beta_n$$

therefore

$$f(\underline{x}) = \sum_{i=n-r+1}^n \alpha_i^2 \lambda_i - \alpha_i \beta_i \quad (\lambda_i \neq 0)$$

therefore

$$\min_{\underline{x}} f(\underline{x}) = \sum_{i=n-r+1}^n \min_{\alpha_i} (\alpha_i^2 \lambda_i - \alpha_i \beta_i) = - \sum \frac{\beta_i^2}{4\lambda_i}.$$

Hence Q has a solution if \underline{b} is orthogonal to every linear combination of $\underline{x}_1, \dots, \underline{x}_{n-r}$.

We conclude this subsection by considering five examples as an illustration of the above.

$$(1) \quad z = y^2 + x^2 = (x \ y) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

The minimum of z is obviously at the origin (see Fig 1).

Note that $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is positive definite and has no zero eigenvalues.

$$(2) \quad z = x^2 + y = (x \ y) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + (x \ y) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

This has no (finite) minimum (see Fig 2). Note that $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

is positive semi-definite and has a zero eigenvalue with eigenvector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

$$(3) \quad z = (x^2 + y^2) + y = (x \ y) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + (x \ y) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Because $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is positive definite, the function,

similar to that of example (2) in other respects, has got a unique minimum (see Fig 3). We know, from the theory of general $f(\underline{x})$, that if $A(= d^2f/d\underline{x}^2)$ is positive definite then there exists a solution.

$$(4) \quad z = \underline{x}^2 - \underline{x} = (\underline{x} \quad \underline{y}) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \underline{x} \\ \underline{y} \end{pmatrix} - (\underline{x} \quad \underline{y}) \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

This is an example of a quadratic function with a positive semidefinite matrix that has got a solution (actually an infinite number of solutions - see Fig 4).

The eigenvector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ corresponding to the zero eigenvalue is orthogonal to $\underline{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

$$(5) \quad z = \underline{y}^2 - \underline{x}^2 = (\underline{x} \quad \underline{y}) \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \underline{x} \\ \underline{y} \end{pmatrix}.$$

We include this example as an illustration of what f may look like when the matrix is non-definite (see Fig 5).

4 THE EQUATION CONSTRAINED PROBLEM

In this section we consider the equation constrained problem

$$E \quad \underset{\underline{x}}{\text{minimize}} \quad f(\underline{x}) \quad \text{subject to} \quad \underline{q}(\underline{x}) = \underline{0}$$

where the vector equation $\underline{q}(\underline{x}) = \underline{0}$ represents m equation constraints of the type $q_i(\underline{x}) = 0$, all of which must be satisfied at the solution \underline{x}^* of E .

Any point \underline{x} which satisfies $\underline{q}(\underline{x}) = \underline{0}$ we shall call *feasible*. Suppose \underline{x} is a feasible point. If $\underline{x} + \Delta\underline{x}$ is also feasible, then $\Delta\underline{x}$ is said to be *feasible at \underline{x}* or sometimes a *feasible direction at \underline{x}* . By a solution of E we mean a feasible point \underline{x}^* such that $f(\underline{x}^*) \leq f(\underline{x}^* + \Delta\underline{x})$ for all small enough feasible directions $\Delta\underline{x}$ at \underline{x}^* . As before we assume that $f(\underline{x})$ is continuously twice differentiable and that $f(\underline{x}^*) > -\infty$. Finally we stress again that our theory concerns *local* solutions to our problems.

We shall begin our discussion by examining the two-dimensional situation and using any insight gained to help us tackle the n -dimensional problem.

4.1 The two-dimensional problem

The simplest equation constrained minimization problem is the two-dimensional

$$E2 \quad \underset{x,y}{\text{minimize}} \quad f(x,y) \quad \text{subject to} \quad q(x,y) = 0$$

The constraint equation can be thought of as a contour in the (x,y) plane. Provided the contour is not everywhere parallel to one of the axes (so this rules out $q(x,y) = x = 0$ for instance) it is possible to rewrite the constraint as

$$y = \psi(x) .$$

The problem E2 becomes

E1 minimize $f(x, \psi(x))$ x

and the solution of this problem is given by

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{dy}{dx} = 0 \quad (4-1-1)$$

where df/dx is merely the gradient of f along the contour $y = \psi(x)$. Suppose, for the moment, that $\frac{\partial f}{\partial y} \neq 0$ (but we shall bear this assumption in mind in the following discussion). (4-1-1) can be rewritten as

$$\frac{dy}{dx} = - \left(\frac{\partial f}{\partial y} \right)^{-1} \frac{\partial f}{\partial x} . \quad (4-1-2)$$

Also we can write the equation constraint as

$$q(x,y) = q(x, \psi(x)) = 0 . \quad (4-1-3)$$

Differentiating (4-1-3) totally with respect to x we get

$$\frac{dq}{dx} = \frac{\partial q}{\partial x} + \frac{\partial q}{\partial y} \frac{dy}{dx} = 0 . \quad (4-1-4)$$

If we assume $\frac{\partial q}{\partial y} \neq 0$, we can write (4-1-4) as

$$\frac{dy}{dx} = - \left(\frac{\partial q}{\partial y} \right)^{-1} \frac{\partial q}{\partial x} . \quad (4-1-5)$$

Equating (4-1-2) and (4-1-5) and rearranging we get

$$\frac{\partial f}{\partial x} \left(\frac{\partial q}{\partial x} \right)^{-1} = \left(\frac{\partial f}{\partial y} \right) \left(\frac{\partial q}{\partial y} \right)^{-1} \quad (4-1-6)$$

where we have assumed $\frac{\partial q}{\partial x} \neq 0$. If we set the value of each side of (4-1-6) equal to $-\lambda$

we can write

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial q}{\partial x} = 0 \quad (4-1-7)$$

$$\frac{\partial f}{\partial y} + \lambda \frac{\partial q}{\partial y} = 0 \quad (4-1-8)$$

Lagrange (1736-1813) noticed that (4-1-7) and (4-1-8) are simply the conditions that are necessarily satisfied by a stationary point of the function

$$\mathcal{L}(x,y,\lambda) = f(x,y) + \lambda q(x,y) \quad (4-1-9)$$

whilst the condition

$$\frac{\partial \mathcal{L}}{\partial \lambda} = q(x,y) = 0 \quad (4-1-10)$$

simply incorporates the constraint into the problem. \mathcal{L} is called the *Lagrangian* or *augmented function* of the problem and λ is called a *Lagrange multiplier*.

We seek some geometrical interpretation of the algebra. Equations (4-1-2) and (4-1-5) simply state that at the solution of E2, the gradients of the contours

$$q(x,y) = 0 \quad (4-1-11)$$

and

$$f(x,y) = f^*$$

are equal, where f^* is the value of f at the solution (x^*, y^*) . We find that this interpretation agrees with our geometrical intuition (see Fig 6). For if the gradients are not parallel, then the contours must intersect at an angle. Except for the special case when the constrained minimum and the unconstrained minimum coincide, this must mean there are points on the constraint contour (4-1-11) on one side or the other of (x^*, y^*) where $f < f^*$, which is a contradiction.

It is important to note that the method of Lagrange multipliers may sometimes be used when the problem functions do not satisfy the assumptions made in the above discussion, namely that $\partial f / \partial y$, $\partial q / \partial x$ and $\partial q / \partial y$ are non-zero. These assumptions were only made in the interest of easing our derivation of (4-1-7) and (4-1-8) and ensured the existence of a *unique* Lagrange multiplier

$$\lambda = - \frac{\partial f}{\partial x} \left(\frac{\partial q}{\partial x} \right)^{-1} = - \frac{\partial f}{\partial y} \left(\frac{\partial q}{\partial y} \right)^{-1}.$$

In fact the method of Lagrange multipliers will work even when our above assumptions do not hold, provided there exists a λ (not necessarily unique) as well as x and y which satisfy (4-1-7), (4-1-8) and (4-1-10). The sufficiency of this is proved in section 4.3 for the more general n dimensional equation constrained case.

We do not bother to prove it here because the reader would gain little geometrical insight from the proof for the present (two-dimensional) case.

4.2 The n -dimensional problem

As usual, we merely extend the notation of the two-dimensional case to the n -dimensional situation. However, we must first make two assumptions about the equation constraints

$$\underline{q}(\underline{x}) = \underline{0} . \quad (4-2-1)$$

The first assumption we make is that the *feasible set* (ie the set of all points satisfying (4-2-1)) is such that there is a path an iterative algorithm can follow from some starting point to the solution \underline{x}^* . For the purposes of these notes, we shall express this succinctly by saying that \underline{x}^* is assumed to be *not isolated*. If no such path exists the point \underline{x}^* is said to be *isolated*. The second assumption we make is much less obvious. We shall assume that

$$\text{rank } \frac{dq}{d\underline{x}}(\underline{x}^*) = m . \quad (4-2-2)$$

In otherwords, the gradient vectors

$$\frac{dq_1}{d\underline{x}}(\underline{x}^*), \dots, \frac{dq_m}{d\underline{x}}(\underline{x}^*)$$

are linearly independent. (4-2-2) is called the *Jacobian assumption*. We make this assumption because (as we shall see) it considerably simplifies the general proof of the method of Lagrange multipliers. However it is important to note that in general the equation constraints will *not* satisfy (4-2-2). The implications of this are discussed more fully in section 4.4.

We wish to solve the problem E where the m equation constraints \underline{q} satisfy the Jacobian assumption. Note that the \underline{q} then also satisfy the conditions of the *implicit function theorem* (see Appendix B). Therefore there exists a vector function $\underline{\psi}$ such that (re-ordering the x_i if necessary)

$$x_i = \psi_i(x_{m+1}, \dots, x_n) \quad i = 1, \dots, m.$$

If we write

$$\underline{u} = (x_{m+1}, \dots, x_n)^T$$

and

$$\underline{v} = (x_1, \dots, x_m)^T$$

then by the Jacobian assumption $\text{rank } (\partial \underline{q} / \partial \underline{v}) = m$ and our problem becomes

$$\text{Eu minimize } f(\underline{\psi}(\underline{u}), \underline{u})$$

From equation (3-2-1) a first-order necessary condition that \underline{u}^* be a solution of Eu is that

$$\frac{df}{d\underline{u}} = \frac{\partial f}{\partial \underline{u}} + \frac{d\underline{v}}{d\underline{u}} \frac{\partial f}{\partial \underline{v}} = \underline{0} . \quad (4-2-3)$$

But

$$\underline{q}(\underline{x}) = \underline{q}(\underline{\psi}(\underline{u}), \underline{u}) = \underline{0} ,$$

therefore

$$\frac{d\underline{q}}{d\underline{u}} = \frac{\partial \underline{q}}{\partial \underline{u}} + \frac{d\underline{v}}{d\underline{u}} \frac{\partial \underline{q}}{\partial \underline{v}} = \underline{0} . \quad (4-2-4)$$

By our above rearrangement, $\partial \underline{q} / \partial \underline{v}$ is non-singular. Hence we can write

$$\frac{d\underline{v}}{d\underline{u}} = - \frac{\partial \underline{q}}{\partial \underline{u}} \left(\frac{\partial \underline{q}}{\partial \underline{v}} \right)^{-1} . \quad (4-2-5)$$

Substituting (4-2-5) into (4-2-3) gives

$$\frac{df}{d\underline{u}} = \frac{\partial f}{\partial \underline{u}} - \frac{\partial \underline{q}}{\partial \underline{u}} \left(\frac{\partial \underline{q}}{\partial \underline{v}} \right)^{-1} \frac{\partial f}{\partial \underline{v}} = \underline{0} . \quad (4-2-6)$$

Now because

$$\left(\frac{\partial \underline{q}}{\partial \underline{v}} \right) \left(\frac{\partial \underline{q}}{\partial \underline{v}} \right)^{-1} = \underline{I}_m \quad (4-2-7)$$

it follows, by postmultiplying (4-2-7) by $\partial f / \partial \underline{v}$ that

$$\frac{\partial \underline{q}}{\partial \underline{v}} \left(\frac{\partial \underline{q}}{\partial \underline{v}} \right)^{-1} \frac{\partial f}{\partial \underline{v}} = \frac{\partial f}{\partial \underline{v}}$$

or rearranging we get

$$\frac{\partial f}{\partial \underline{v}} - \frac{\partial \underline{q}}{\partial \underline{v}} \left(\frac{\partial \underline{q}}{\partial \underline{v}} \right)^{-1} \frac{\partial f}{\partial \underline{v}} = \underline{0} . \quad (4-2-8)$$

If we write

$$\underline{\lambda} = - \left(\frac{\partial \underline{q}}{\partial \underline{v}} \right)^{-1} \frac{\partial f}{\partial \underline{v}} \quad (4-2-9)$$

then equations (4-2-6) and (4-2-8) can be written as

$$\frac{\partial f}{\partial \underline{u}} + \frac{\partial \underline{q}}{\partial \underline{u}} \underline{\lambda} = \underline{0} \quad (4-2-10)$$

and

$$\frac{\partial f}{\partial \underline{v}} + \frac{\partial \underline{q}}{\partial \underline{v}} \underline{\lambda} = \underline{0} . \quad (4-2-11)$$

But \underline{u} and \underline{v} are merely partitions of a rearrangement of our original vector \underline{x} , so that by recombining \underline{u} and \underline{v} , and rearranging if necessary, (4-2-10) and (4-2-11) can be written succinctly as

$$\frac{\partial f}{\partial \underline{x}} + \frac{\partial q}{\partial \underline{x}} \underline{\lambda} = \underline{0}.$$

Thus a first order necessary condition that \underline{x}^* be a solution of E is that

$$\frac{\partial \mathcal{L}}{\partial \underline{x}} (\underline{x}^*) = \underline{0} \quad (4-2-12)$$

where $\mathcal{L}(\underline{x}) = f(\underline{x}) + \underline{\lambda}^T \underline{q}$. As already discussed for the two-dimensional case, the condition $\frac{\partial \mathcal{L}}{\partial \underline{\lambda}} (\underline{x}^*, \underline{\lambda}) = \underline{0}$ is just a restatement of the constraints $\underline{q}(\underline{x}) = \underline{0}$. Notice also that because of (4-2-9), the Jacobian assumption guarantees the existence of *unique* Lagrange multipliers $\underline{\lambda}$.

In view of (4-2-12) therefore, the solution of any equation constrained problem E , satisfying the Jacobian assumption, is also a stationary point of the associated Lagrangian function \mathcal{L} . This is a very useful result because algorithms to find a stationary point of the equivalent Lagrangian problem are of course much easier to design than algorithms to solve the original problem E .

4.3 Second order conditions

For the reader's convenience we begin this section by stating the second order necessary and sufficient conditions that \underline{x}^* be a (local) solution of E . As we shall see, they are very easy to prove.

The *second order necessary condition* is that

$$\underline{\Delta x}^T \frac{d^2 \mathcal{L}}{d \underline{x}^2} (\underline{x}^*) \underline{\Delta x} \geq 0 \quad (4-3-1)$$

for all small enough vectors $\underline{\Delta x}$ that satisfy $\underline{\Delta x}^T \frac{dq}{d \underline{x}} (\underline{x}^*) = \underline{0}^T$. The sufficient conditions are

$$\left. \begin{aligned} \frac{\partial \mathcal{L}}{\partial \underline{x}} (\underline{x}^*) &= \underline{0} \\ \frac{\partial \mathcal{L}}{\partial \underline{\lambda}} (\underline{x}^*) &= \underline{q}(\underline{x}^*) = \underline{0} \\ \underline{\Delta x}^T \frac{d^2 \mathcal{L}}{d \underline{x}^2} (\underline{x}^*) \underline{\Delta x} &> 0 \end{aligned} \right\} \quad (4-3-2)$$

and

for all vectors $\underline{\Delta x}$ that satisfy $\underline{\Delta x}^T \frac{dq}{d \underline{x}} (\underline{x}^*) = \underline{0}^T$.

It is left to the reader to provide the parallel discussion concerning the two-dimensional case if he still requires geometrical insight. The similarities in notation are such that he should have no difficulty.

In section 4.2 we only showed that \underline{x}^* is a stationary point of f . However, it is possible that there are other stationary points as well. Thus an algorithm designed to converge at a stationary point of f may not converge at \underline{x}^* . In view of the sufficiency conditions (4-3-2), we see that if the algorithm is designed to converge at a stationary point \underline{x}^* of f such that (4-3-2) is also satisfied then \underline{x}^* is a solution of F . It is for this reason that the sufficiency conditions are important.

We derive (4-3-3) as follows. As in section 4.2, since the constraints q_i satisfy the Jacobian assumption, our problem F is equivalent to the unconstrained problem F_u . From (4-2-2), a second order necessary condition that \underline{u}^* be a solution of F_u is that

$$\underline{\Delta u}^T \frac{d^2 f}{d\underline{u}^2} \left(\underline{y}(\underline{u}^*), \underline{u}^* \right) \underline{\Delta u} > 0 \quad (4-3-3)$$

for all vectors $\underline{\Delta u}$. From (A-1) we have that

$$\frac{d^2 f}{d\underline{u}^2} = \frac{d\underline{x}}{d\underline{u}} \frac{d^2 f}{d\underline{x}^2} \left(\frac{d\underline{x}}{d\underline{u}} \right)^T + \sum_j \frac{\partial f}{\partial v_j} \frac{d^2 v_j}{d\underline{u}^2} \quad (4-3-4)$$

where the total derivative is taken along the contour $\underline{u} = \underline{u}(\underline{v})$. But $\lambda_i q_i(\underline{y}(\underline{u}), \underline{u}) = 0$ for $i = 1, \dots, m$ and for all \underline{u} .

Differentiating twice we have, again from (A-1), that

$$\frac{d^2}{d\underline{u}^2} (\lambda_i q_i) = \frac{d\underline{x}}{d\underline{u}} \frac{d^2}{d\underline{x}^2} (\lambda_i q_i) \left(\frac{d\underline{x}}{d\underline{u}} \right)^T + \sum_j \frac{\partial q_i}{\partial v_j} \frac{d^2 v_j}{d\underline{u}^2} \quad (4-3-5)$$

for $i = 1, \dots, m$. Adding (4-3-4) and (4-3-5) we have

$$\frac{d^2 f}{d\underline{u}^2} = \frac{d^2}{d\underline{u}^2} \left[f + \sum_j \lambda_j q_j \right] = \frac{d\underline{x}}{d\underline{u}} \left[\frac{d^2 f}{d\underline{x}^2} + \sum_i \lambda_i \frac{d^2 q_i}{d\underline{x}^2} \right] \left(\frac{d\underline{x}}{d\underline{u}} \right)^T + \sum_j \frac{\partial}{\partial v_j} \left(f + \sum_i \lambda_i q_i \right) \frac{d^2 v_j}{d\underline{u}^2}.$$

But $\frac{df}{d\underline{x}}(\underline{x}^*) = 0$ means that in particular $\frac{\partial}{\partial v_j} (f + \sum_i \lambda_i q_i) = 0$ for $j = 1, \dots, m$.

Therefore

$$\frac{d^2 f}{d\underline{u}^2} = \frac{d\underline{x}}{d\underline{u}} \frac{d^2 f}{d\underline{x}^2} \left(\frac{d\underline{x}}{d\underline{u}} \right)^T \quad (4-3-6)$$

But along the contour $\underline{u} = \underline{u}(\underline{v})$, $d^2\mathcal{L}/d\underline{u}^2 = d^2f/d\underline{u}^2$. In view of (4-3-3), equation (4-3-6) means that

$$\Delta \underline{u}^T \frac{d\underline{x}}{d\underline{u}} \frac{d^2\mathcal{L}}{d\underline{x}^2} \left(\frac{d\underline{x}}{d\underline{u}} \right)^T \Delta \underline{u} \geq 0 \quad (4-3-7)$$

for all vectors $\Delta \underline{u}$.

Let $\Delta \underline{x}$ be any vector satisfying

$$\Delta \underline{x}^T \frac{d\underline{q}}{d\underline{x}} = \underline{0}^T. \quad (4-3-8)$$

Partitioning $\Delta \underline{x}^T$ into $(\Delta \underline{v}^T : \Delta \underline{u}^T)$ where $\Delta \underline{v}$ and $\Delta \underline{u}$ are column vectors of appropriate length, (4-3-8) can be rewritten as

$$\Delta \underline{v}^T \frac{\partial \underline{q}}{\partial \underline{v}} + \Delta \underline{u}^T \frac{\partial \underline{q}}{\partial \underline{u}} = \underline{0}^T.$$

By the Jacobian assumption, $\partial \underline{q}/\partial \underline{v}$ is nonsingular.

Hence postmultiplying by $\left(\frac{\partial \underline{q}}{\partial \underline{v}} \right)^{-1}$ gives

$$\Delta \underline{v}^T = - \Delta \underline{u}^T \frac{\partial \underline{q}}{\partial \underline{u}} \left(\frac{\partial \underline{q}}{\partial \underline{v}} \right)^{-1}.$$

In view of (4-2-5) we have shown that

$$\Delta \underline{v}^T = \Delta \underline{u}^T \frac{d\underline{v}}{d\underline{u}}.$$

Hence

$$\Delta \underline{x}^T = \Delta \underline{u}^T \left(\frac{d\underline{v}}{d\underline{u}} : \underline{I} \right) = \Delta \underline{u}^T \frac{d\underline{x}}{d\underline{u}}. \quad (4-3-9)$$

We have, therefore, proved the second order necessary condition since in view of (4-3-7), equation (4-3-9) implies that

$$\Delta \underline{x}^T \frac{d^2\mathcal{L}}{d\underline{x}^2} \Delta \underline{x} \geq 0$$

for any vector $\Delta \underline{x}$ satisfying (4-3-8).

We now prove the sufficiency conditions (4-3-2). Note carefully that they imply the existence of a set of (not necessarily unique) Lagrange multipliers $\underline{\lambda}$. For reasons discussed in the next subsection, we prove the sufficiency of (4-3-2) *without* appealing to the Jacobian assumption. We shall assume that (4-3-2) holds and that \underline{x}^* is not a solution of E . We then obtain a proof by arriving at a contradiction.

If \underline{x}^* is a feasible point and not a solution of E , then since \underline{x}^* is not isolated (see section 4.2), there must exist a sequence of feasible points $\underline{x}_n = \underline{x}^* + \Delta \underline{x}_n$ which converge to \underline{x}^* and which satisfy $f(\underline{x}_n) \leq f(\underline{x}^*)$, $n = 1, 2, \dots$. Let $\hat{\underline{u}}_n$ be the unit vector along the direction of $\Delta \underline{x}_n$. Then the sequence $\hat{\underline{u}}_1, \hat{\underline{u}}_2, \dots$ is obviously bounded and $\Delta \underline{x}_n = \Delta x_n \hat{\underline{u}}_n$ where $\Delta x_n \geq 0$ is the magnitude of $\Delta \underline{x}_n$. It is easy to show that $\Delta x_n \rightarrow 0$ as $n \rightarrow \infty$.

It is well-known that a bounded sequence has a convergent subsequence⁵. Hence $\hat{\underline{u}}_1, \hat{\underline{u}}_2, \dots$ has a convergent subsequence. Let this subsequence be $\underline{y}_1, \underline{y}_2, \dots$ and suppose it converges to \underline{y} . Let the corresponding modulus of each vector \underline{y}_n be Δy_n such that $\underline{z}_n = \underline{x}^* + \Delta y_n \underline{y}_n$ is a feasible vector.

Now since each \underline{z}_n is feasible we have that

$$q(\underline{z}_n) - q(\underline{x}^*) = 0.$$

Dividing by Δy_n , we have

$$\frac{q(\underline{x}^* + \Delta y_n \underline{y}_n) - q(\underline{x}^*)}{\Delta y_n} = 0.$$

Now from Taylor's theorem for vector functions of vector variables (see Appendix B) we have that

$$\frac{q(\underline{x}^* + \Delta y_n \underline{y}_n) - q(\underline{x}^*)}{\Delta y_n} = \left(\frac{dq}{d\underline{x}}(\underline{x}^*) \right)^T \underline{y}_n + o(\Delta y_n) = 0$$

where we interpret the symbol $o(\Delta y_n)$ as a vector of symbols $o(\Delta y_n)$. On letting $n \rightarrow \infty$ we see that

$$\underline{y}^T \frac{dq}{d\underline{x}}(\underline{x}^*) = 0^T.$$

Let λ_i be the Lagrange multiplier corresponding to the constraint $q_i(\underline{x})$. Then from Taylor's theorem for scalar functions we have for $j = 1, \dots, m$ that

$$0 = \lambda_i q_i(\underline{y}_n) = \lambda_i q_i(\underline{x}^*) + \lambda_i \Delta y_n \underline{y}_n^T \frac{dq_i}{d\underline{x}}(\underline{x}^*) + \frac{1}{2} \Delta y_n^2 \underline{y}_n^T \frac{d^2 q_i}{d\underline{x}^2}(\underline{x}^*) \underline{y}_n + o(\Delta y_n^3). \quad (4-3-10)$$

Also

$$0 \geq f(\underline{z}_n) - f(\underline{x}^*) = \Delta y_n \underline{y}_n^T \frac{df}{d\underline{x}}(\underline{x}^*) + \frac{1}{2} \Delta y_n^2 \underline{y}_n^T \frac{d^2 f}{d\underline{x}^2}(\underline{x}^*) \underline{y}_n + o(\Delta y_n^3). \quad (4-3-11)$$

Adding (4-3-10) for $j = 1, \dots, m$ to (4-3-11) we obtain

$$0 \geq \Delta y_n^T \frac{df}{dx}(\underline{x}^*) + \frac{1}{2} \Delta y_n^T \frac{d^2f}{dx^2} \underline{y}_n + o(\Delta y_n^3). \quad (4-3-12)$$

We are now in a position to obtain our contradiction.

If the sufficient conditions (4-3-2) hold then $\frac{df}{dx}(\underline{x}^*) = \underline{0}$.

Therefore (4-3-12) becomes

$$0 \geq \frac{1}{2} \Delta y_n^T \frac{d^2f}{dx^2} \underline{y}_n + o(\Delta y_n^3).$$

Multiplying by $2/\Delta y_n^2$ we get

$$0 \geq \underline{y}_n^T \frac{d^2f}{dx^2} \underline{y}_n + o(\Delta y_n).$$

Letting $n \rightarrow \infty$ we see that

$$0 \geq \underline{y}^T \frac{d^2f}{dx^2} \underline{y}. \quad (4-3-13)$$

But we have shown that \underline{y} satisfies $\underline{y}^T \frac{dq}{dx}(\underline{x}^*) = \underline{0}^T$. Therefore (4-3-13) contradicts the sufficient condition and we have finished our proof.

4.4 Implications of the Jacobian assumption

There appears to be very little discussion of the Jacobian assumption in the literature (but see Fiacco and McCormick⁶). Almost always the assumption is made without any comment or qualification. More importantly, from the beginner's point of view, it is made without motivation. But this motivation is simply that, as we have seen, the assumption guarantees, because of (4-2-9), that there exist *unique* Lagrange multipliers $\underline{\lambda}$ such that

$$\frac{df}{dx}(\underline{x}^*) + \frac{dq}{dx}(\underline{x}^*) \underline{\lambda} = \underline{0}. \quad (4-4-1)$$

Since practical problems need not satisfy the Jacobian assumption it seems desirable to explore the consequences of removing it. If we do so, then we can no longer be sure that any $\underline{\lambda}$ exist (even non-uniquely) such that (4-4-1) holds. For instance, consider the problem

$$\text{minimize } f(x, y, z) = x^2 + (y - 1)^2 + (z + 1)^2$$

$$\text{subject to } q_1(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$$

$$\text{and } q_2(x, y, z) = x^2 + (y - 2)^2 + z^2 - 1 = 0.$$

It is clear (see Fig 7) that the only feasible point is $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ which therefore must be the

solution of the problem. Define the Lagrangian by $\mathcal{L} = f + \lambda_1 q_1 + \lambda_2 q_2$, then for this example (4-2-12) becomes

$$\frac{\partial \mathcal{L}}{\partial x} = 2x + 2x\lambda_1 + 2x\lambda_2 = 0$$

$$\frac{\partial \mathcal{L}}{\partial y} = 2y + 2y\lambda_1 + 2(y-2)\lambda_2 = 0$$

$$\frac{\partial \mathcal{L}}{\partial z} = 2(z+1) + 2z\lambda_1 + 2z\lambda_2 = 0 \quad (4-4-2)$$

No values of λ_1 and λ_2 can satisfy (4-4-2) at $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. We have included this example

to illustrate the important fact that even problems with continuous functions f and q can have isolated solutions. However, if we assume the solution of our equation constrained problem E is not isolated and that a further assumption (discussed in the proof below) also holds then we can show that Lagrange multipliers *must* exist (if not uniquely).

As in the proof of the sufficiency of (4-3-2) we can construct a sequence of feasible points $\underline{z}_n = \underline{x}^* + \Delta y_n \underline{y}_n$ which converges to \underline{x}^* , where \underline{y}_n converges to a unit vector \underline{y} . We have already shown that

$$\underline{y}^T \frac{dq}{d\underline{x}}(\underline{x}^*) = 0^T \quad (4-4-3)$$

which means that \underline{y} is orthogonal to any linear combination of the gradient vectors $\frac{dq_i}{d\underline{x}}(\underline{x}^*)$. Note that in all sequences of feasible vectors of the form $\underline{z}_n = \underline{x}^* + \Delta y_n \underline{y}_n$ which converge to \underline{x}^* , the \underline{y}_n will converge to some vector \underline{y} which satisfies (4-4-3).

Let r be the number of gradient vectors $\frac{dq_i}{d\underline{x}}(\underline{x}^*)$ that are linearly independent. By renumbering the $\frac{dq_i}{d\underline{x}}(\underline{x}^*)$ if necessary we can assume that $\frac{dq_1}{d\underline{x}}(\underline{x}^*), \dots, \frac{dq_r}{d\underline{x}}(\underline{x}^*)$ are linearly independent. It is easy to show that there exist exactly $n - r$ orthonormal vectors \underline{v}_j , say, which in addition are orthogonal to any linear combination of the $\frac{dq_i}{d\underline{x}}(\underline{x}^*)$. We now prove that $\frac{df}{d\underline{x}}(\underline{x}^*)$ is also orthogonal to each of these \underline{v}_j . But to prove this we need an additional assumption (as mentioned above).

For each \underline{v}_j we construct a sequence of the form

$$\underline{z}_n = \underline{x}^* + \Delta y_n (\underline{v}_j + \underline{y}_n) \quad (4-4-4)$$

where Δy_n is a sequence of positive numbers converging to 0. Our additional assumption is that we can choose the vectors \underline{y}_n to be a sequence converging to $\underline{0}$ such that each vector \underline{z}_n is a feasible point. Of course, it is sometimes *not* possible to choose such vectors. Consider the problem

$$\text{minimize } f = x^2 + (y - 1)^2 + (z + 1)^2$$

$$\text{subject to } q_1 = x^2 + y^2 - 1 = 0$$

$$\text{and } q_2 = x^2 + y^2 + z^2 - 1 = 0.$$

The feasible set is the circle

$$\left. \begin{aligned} x^2 + y^2 &= 1 \\ z &= 0 \end{aligned} \right\}$$

and the solution of the problem is easily seen to be $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ where the gradient vectors of

both q_1 and q_2 are both $\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$ whilst the gradient vector of f is $\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$.

Now the two orthonormal vectors that are orthogonal to $\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$ are in this case

clearly $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. In particular, for $\underline{v}_j = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ we see from Fig 8 that for \underline{z}_n

to be feasible the vector \underline{y}_n has to have a length of at least one unit and hence cannot converge to $\underline{0}$.

We now return to our proof and assume that the vectors \underline{y}_n exist as required in (4-4-4). Since \underline{z}_n is bounded, \underline{z}_n (or a subsequence) must converge, and the limit point is clearly \underline{x}^* .

Now by Taylor's theorem

$$f(\underline{z}_n) = f(\underline{x}^*) + \Delta y_n (\underline{v}_j + \underline{y}_n)^T \frac{df}{d\underline{x}}(\underline{x}^*) + O(\Delta y_n^2). \quad (4-4-5)$$

But since $f(\underline{x}^*)$ is a minimum of f we have $f(\underline{z}_n) \geq f(\underline{x}^*)$. In view of (4-4-5), this means that

$$f(\underline{x}^*) + \Delta y_n (\underline{v}_j + \underline{y}_n)^T \frac{df}{d\underline{x}}(\underline{x}^*) + 0(\Delta y_n^2) \geq f(\underline{x}^*) .$$

Subtracting $f(\underline{x}^*)$ from both sides and dividing by $\Delta y_n > 0$ we obtain

$$(\underline{v}_j + \underline{y}_n)^T \frac{df}{d\underline{x}}(\underline{x}^*) + 0(\Delta y_n) \geq 0 . \quad (4-4-6)$$

Letting $n \rightarrow \infty$, we find that $\underline{y}_n \rightarrow 0$ and $\Delta y_n \rightarrow 0+$ and we see that (4-4-6) implies

$$\underline{v}_j^T \frac{df}{d\underline{x}}(\underline{x}^*) \geq 0 . \quad (4-4-7)$$

But we could have equally well chosen Δy_n in (4-4-4) to be a sequence of *negative* numbers. Then division by Δy_n reverses the inequality sign in the calculation above and instead of (4-4-7) we obtain

$$\underline{v}_j^T \frac{df}{d\underline{x}}(\underline{x}^*) \leq 0 \quad (4-4-8)$$

(4-4-7) and (4-4-8) together imply that $\underline{v}_j^T \frac{df}{d\underline{x}}(\underline{x}^*) = 0$. Hence we have shown that

$\frac{df}{d\underline{x}}(\underline{x}^*)$ is orthogonal to each of the \underline{v}_j .

Now the vectors $\underline{v}_1, \dots, \underline{v}_{n-r}$ together with the r linearly independent vectors $\frac{dq_i}{d\underline{x}}(\underline{x}^*)$ form a basis. This means that any vector, and in particular the vector $\frac{df}{d\underline{x}}(\underline{x}^*)$ is a linear combination of the \underline{v}_j and the $\frac{dq_i}{d\underline{x}}(\underline{x}^*)$. That is, there exist numbers $\alpha_1, \dots, \alpha_{n-r}$ and $\lambda_1, \dots, \lambda_r$ such that

$$\frac{df}{d\underline{x}}(\underline{x}^*) = \alpha_1 \underline{v}_1 + \dots + \alpha_{n-r} \underline{v}_{n-r} + \lambda_1 \frac{dq_1}{d\underline{x}}(\underline{x}^*) + \dots + \lambda_r \frac{dq_r}{d\underline{x}}(\underline{x}^*) . \quad (4-4-9)$$

But, since the \underline{v}_j are orthonormal and are orthogonal to $\frac{df}{d\underline{x}}(\underline{x}^*)$ and every $\frac{dq_i}{d\underline{x}}(\underline{x}^*)$, by premultiplying (4-4-9) by \underline{v}_j^T ($j = 1, \dots, n-r$) it is easy to see that $\alpha_j = 0$, ($j = 1, \dots, n-r$).

If we also put $\lambda_{r+1} = \dots = \lambda_n = 0$, we see that we have established the existence of a set of numbers λ_i satisfying (4-4-1). Thus the λ_i are Lagrange multipliers and we have finished our proof.

It follows that, in general, practical problems *will* have $\underline{\lambda}$ satisfying (4-4-1). Hence we have also proved that, in general, an algorithm that solves the equation constrained problem E by finding a $\underline{\lambda}$ to satisfy the sufficient conditions (4-3-2) will

be successful. The corollary is, of course, that such an algorithm must have in-built safeguards to prevent it from giving misleading results in cases (such as our example) where no Lagrange multipliers exist.

5 THE KUHN-TUCKER CONDITIONS

In this section we shall consider the general constrained optimization problem G , where there are constraints of both equation and inequality type.

$$\begin{array}{ll} G & \text{minimize } f(\underline{x}) \text{ subject to } \underline{q}(\underline{x}) = \underline{0} \\ & \underline{x} \quad \text{and } \underline{c}(\underline{x}) \leq \underline{0} \end{array}$$

where $\underline{c}(\underline{x})$ is an $m' \times 1$ column vector valued function of \underline{x} .

We have postponed the derivation of the necessary and sufficient conditions of this problem until after discussing the equation constrained problem E . The reader may feel that this is because the conditions for the problem G are more difficult to derive than for any of the problems mentioned earlier. This is not so, for, as we shall see, the general constrained problem G can be quite readily transformed into the problem E . Hence if we can derive necessary and sufficient conditions for E , we can also do so for G . It is for this reason that we have left our discussion until this stage. As always we shall best proceed by considering related but simpler problems.

5.1 A one-dimensional problem

We consider first the problem

$$Z1 \quad \text{minimize } f(x) \text{ subject to } x \geq 0$$

To derive a first order necessary condition, we proceed as for the unconstrained problem $U1$, by expanding $f(x)$ by Taylor's series about a local solution x^* ,

$$f(x^* + \Delta x) = f(x^*) + \Delta x \frac{df}{dx}(x^*) + O(\Delta x^2) \quad (5-1-1)$$

Now x^* is a local solution of $Z1$ means that

$$f(x^* + \Delta x) \geq f(x^*) \quad (5-1-2)$$

for all small enough Δx satisfying the constraint $x^* + \Delta x \geq 0$. Using (5-1-1) to eliminate $f(x^* + \Delta x)$ from (5-1-2) we obtain

$$f(x^*) + \Delta x \frac{df}{dx}(x^*) + O(\Delta x^2) \geq f(x^*)$$

or

$$\Delta x \frac{df}{dx}(x^*) + O(\Delta x^2) \geq 0 \quad (5-1-3)$$

Note that for x^* to be a solution of Z1, also implies that x^* satisfies the constraint, i.e. that $x^* \geq 0$. If $x^* = 0$, then $x^* + \Delta x$ satisfies the constraint only if $\Delta x \geq 0$. Dividing (5-1-3) by $\Delta x > 0$ and letting $\Delta x \rightarrow 0^+$, we obtain

$$\frac{df}{dx}(x^*) \geq 0 \quad (5-1-4)$$

but if $x^* > 0$, then $x^* + \Delta x$ satisfies the constraint if $\Delta x \geq -x^*$. Dividing (5-1-3) by such $\Delta x < 0$ and letting $\Delta x \rightarrow 0^-$ we get

$$\frac{df}{dx}(x^*) \leq 0. \quad (5-1-5)$$

If $x^* > 0$, $x^* + \Delta x$ will also satisfy the constraint if $\Delta x \geq 0 > -x^*$. Hence (5-1-4) will also. Since (5-1-4) and (5-1-5) both hold, we have that

$$\frac{df}{dx}(x^*) = 0. \quad (5-1-6)$$

We summarize the above. A necessary condition that x^* be a solution of Z1 is that

$$\frac{df}{dx}(x^*) \leq 0 \quad \text{if } x^* = 0 \quad (5-1-7)$$

and

$$\frac{df}{dx}(x^*) = 0 \quad \text{if } x^* > 0. \quad (5-1-8)$$

It is customary to abbreviate (5-1-7) and (5-1-8) into the one condition

$$x^* \frac{df}{dx}(x^*) = 0. \quad (5-1-9)$$

Since x^* must also satisfy the constraint $x^* \geq 0$, we readily see that (5-1-9) is in fact equivalent to (5-1-7) and (5-1-8).

The geometrical interpretation of (5-1-9) is straightforward, though perhaps not obvious to the beginner. (5-1-9) simply states that x^* is either on the constraint (i.e. $x^* = 0$) or it is not. If x^* is not on the constraint, then the constraint in no way restricts x^* and hence our problem Z1 is equivalent to the unconstrained problem U1. If x^* does lie on the constraint boundary, then we have solved the problem Z1 and $x^* = 0$.

5.2 A special n-dimensional problem

We can readily extend the method of section 5.1 to deal with the problem.

$\begin{array}{l} Z \\ \underline{x} \end{array} \text{ minimize } f(\underline{x}) \text{ subject to } \underline{x} \geq \underline{0}$

We expand $f(\underline{x})$ about a local solution \underline{x}^*

$$f(\underline{x}^* + \underline{\Delta x}) = f(\underline{x}^*) + \underline{\Delta x}^T \frac{df}{d\underline{x}}(\underline{x}^*) + O|\underline{\Delta x}|^2.$$

Now \underline{x}^* is a local solution of Z so that $f(\underline{x}^* + \underline{\Delta x}) \geq f(\underline{x}^*)$ for all small enough feasible $\underline{\Delta x}$ at \underline{x}^* . In this case $\underline{\Delta x}$ is feasible means that $\underline{x}^* + \underline{\Delta x} \geq \underline{0}$. We obtain in the usual manner that

$$\underline{\Delta x}^T \frac{df}{d\underline{x}}(\underline{x}^*) + O|\underline{\Delta x}|^2 \geq 0. \quad (5-2-1)$$

In particular, (5-2-1) holds for $\underline{\Delta x} = \Delta x \hat{e}_i$ where \hat{e}_i is the i th unit co-ordinate vector and Δx is the magnitude of $\underline{\Delta x}$.

Hence

$$\Delta x_i \frac{\partial f}{\partial x_i}(\underline{x}^*) + O(\Delta x_i)^2 \geq 0. \quad (5-2-2)$$

Now (5-2-2) is analogous to (5-1-3) and we can follow exactly the same procedure as in section 5.1 to obtain results equivalent to (5-1-7) and (5-1-8), namely

$$\frac{\partial f}{\partial x_i}(\underline{x}^*) \geq 0 \quad \text{if } x_i^* = 0 \quad (5-2-3)$$

and

$$\frac{\partial f}{\partial x_i}(\underline{x}^*) = 0 \quad \text{if } x_i^* > 0. \quad (5-2-4)$$

As in section 5.1, we abbreviate these to

$$x_i^* \frac{\partial f}{\partial x_i}(\underline{x}^*) = 0. \quad (5-2-5)$$

We can repeat the same procedure for all i . Adding the n equations of the form (5-2-5) together we obtain

$$\underline{x}^{*T} \frac{df}{d\underline{x}}(\underline{x}^*) = 0. \quad (5-2-6)$$

Note that this condition is in fact equivalent to (5-2-5) because \underline{x}^* satisfies the conditions (5-2-3) and (5-2-4). Hence (5-2-6) is also equivalent to (5-2-3) and (5-2-4) and is thus a first order necessary condition that \underline{x}^* be a solution of Z .

5.3 Active constraints and the Jacobian assumption

Let \underline{x}^* be a solution of the general constrained problem G . Then for each inequality constraint $c_i(\underline{x}) \leq 0$, \underline{x}^* either lies on the constraint boundary

(ie $c_i(\underline{x}^*) = 0$) or it does not. If \underline{x}^* does lie on the boundary, the constraint is said to be *active*.

By rearranging the inequality constraints if necessary, we can partition the vector $\underline{c}^T(\underline{x})$ into $(\underline{c}_A^T(\underline{x}) : \underline{c}_B^T(\underline{x}))$ where $\underline{c}_A(\underline{x}) \leq 0$ denotes the active constraints and $\underline{c}_B(\underline{x}) \leq 0$ denotes the not active constraints.

As we did for the equation constrained problem, to ensure the existence of unique Lagrange multipliers we shall have to assume that the columns of

$$\left(\frac{dq}{d\underline{x}}(\underline{x}^*) : \frac{dc_A}{d\underline{x}}(\underline{x}^*) \right)$$

are linearly independent. This is the Jacobian assumption for the general constrained problem G.

5.4 The general constrained problem

The first order necessary conditions that \underline{x}^* be a local solution of G are as follows. There exists an $m' \times 1$ vector $\underline{\mu}$ and an $m \times 1$ vector $\underline{\lambda}$ such that

$$\left. \begin{aligned} \frac{df}{d\underline{x}}(\underline{x}^*) + \frac{dq}{d\underline{x}}(\underline{x}^*)\underline{\lambda} + \frac{dc}{d\underline{x}}(\underline{x}^*)\underline{\mu} &= 0 \\ \underline{\mu}^T \underline{c}(\underline{x}^*) &= 0 \\ \underline{\mu} &\geq 0 \end{aligned} \right\} \quad (5-4-1)$$

and

(5-4-1) are called the *Kuhn-Tucker conditions*. The vectors $\underline{\lambda}$ and $\underline{\mu}$ are called Lagrange multipliers. (The $\underline{\mu}$ are sometimes called *Kuhn-Tucker multipliers* to emphasise their being distinct from the $\underline{\lambda}$.) We can immediately derive (5-4-1) from the results we have obtained earlier.

Now \underline{x}^* is also a solution of the problem

$$\underset{\underline{x}}{\text{C minimize}} \quad f(\underline{x}) \quad \text{subject to} \quad q(\underline{x}) = 0 \quad \text{and} \quad \underline{c}_A(\underline{x}) = 0$$

at least in a small enough region around \underline{x}^* . For if \underline{x}^* is a local solution of G then

$$f(\underline{x}^*) \leq f(\underline{x}^* + \underline{\Delta x}) \quad (5-4-2)$$

for all small enough $\underline{\Delta x}$ such that $q(\underline{x}^* + \underline{\Delta x}) = 0$ and $\underline{c}(\underline{x}^* + \underline{\Delta x}) \leq 0$. Now since $\underline{c}(\underline{x})$ is continuous and $\underline{c}_B(\underline{x}^*) < 0$, then by the intermediate value theorem (see Appendix B) $\underline{c}_B(\underline{x}^* + \underline{\Delta x}) < 0$ for all small enough $\underline{\Delta x}$. Hence, provided $\underline{\Delta x}$ is small enough, $\underline{x}^* + \underline{\Delta x}$ will automatically satisfy the non-active constraints. Hence (5-4-2) will hold for all small enough $\underline{\Delta x}$ such that $q(\underline{x}^* + \underline{\Delta x}) = 0$ and $\underline{c}_A(\underline{x}^* + \underline{\Delta x}) \leq 0$.

In particular, (5-4-2) must hold for all small enough $\Delta \underline{x}$ such that $\underline{c}_A(\underline{x}^* + \Delta \underline{x}) = \underline{0}$. Hence \underline{x}^* is also a solution of C.

But C is an equation constrained problem, whose first order necessary conditions are given by (4-2-12). The Lagrangian for problem C can be written

$$\mathcal{L}(\underline{x}) = f(\underline{x}) + \underline{\lambda}^T \underline{q}(\underline{x}) + \underline{v}^T \underline{c}_A(\underline{x}),$$

where \underline{v} is a column vector of appropriate length.

Hence if \underline{x}^* is a solution of C then

$$\frac{df}{d\underline{x}}(\underline{x}^*) + \frac{dq}{d\underline{x}}(\underline{x}^*) \underline{\lambda} + \frac{dc_A}{d\underline{x}}(\underline{x}^*) \underline{v} = \underline{0}$$

or

$$\frac{df}{d\underline{x}}(\underline{x}^*) = - \frac{dq}{d\underline{x}}(\underline{x}^*) \underline{\lambda} - \frac{dc_A}{d\underline{x}}(\underline{x}^*) \underline{v}. \quad (5-4-3)$$

We now show that $\underline{v} \geq \underline{0}$. Suppose instead that $v_j < 0$ for at least one j . Let $d\underline{x}$ be a feasible vector at \underline{x}^* . Note that we can choose $d\underline{x}$ such that

$$c_i(\underline{x}^* + d\underline{x}) = 0 \quad i \neq j$$

$$c_j(\underline{x}^* + d\underline{x}) < 0$$

since otherwise $c_j(\underline{x})$ is functionally dependent on the other $c_i(\underline{x})$. Thus $c_j = \phi(\underline{c}_A)$ say. Differentiating with respect to \underline{x} gives

$$\frac{dc_j}{d\underline{x}} = \frac{dc_A}{d\underline{x}} \frac{d\phi}{dc_A}.$$

We see that $dc_j/d\underline{x}$ is a linear combination of the other constraint gradients. In particular, at \underline{x}^* this contradicts the Jacobian assumption. Now

$$0 = [q_i(\underline{x}^* + d\underline{x}) - q_i(\underline{x}^*)] \lambda_i = d\underline{x}^T \frac{dq_i}{d\underline{x}} \lambda_i + 0|d\underline{x}|^2$$

for $i = 1, \dots, m$. Thus

$$- d\underline{x}^T \frac{dq_i}{d\underline{x}} \lambda_i = 0|d\underline{x}|^2.$$

Similarly

$$- d\underline{x}^T \frac{dc_i}{d\underline{x}} v_i = 0|d\underline{x}|^2 \quad \text{for } i \neq j.$$

Now from (5-4-3) we have that

$$\begin{aligned} d\mathbf{x}^T \frac{df}{d\mathbf{x}} &= - \sum_i d\mathbf{x}^T \frac{dq_i}{d\mathbf{x}} \lambda_i - \sum_{i \neq j} d\mathbf{x}^T \frac{dc_i}{d\mathbf{x}} v_i - d\mathbf{x}^T \frac{dc_j}{d\mathbf{x}} v_j + 0|d\mathbf{x}|^2 \\ &= - d\mathbf{x}^T \frac{dc_j}{d\mathbf{x}} v_j + 0|d\mathbf{x}|^2 . \end{aligned}$$

This can be written as

$$df = - dc_j v_j + 0|d\mathbf{x}|^2 .$$

But since $d\mathbf{x}$ is feasible we must have that $dc_j < 0$. Also $v_j < 0$, therefore $k > 0$ where $k = - dc_j v_j$. Thus

$$df = k + 0|d\mathbf{x}|^2 .$$

Since k is of order $|d\mathbf{x}|$, and the functions of interest are continuous, we can find $d\mathbf{x}$ small enough so that $df < 0$, which contradicts our assumption that \mathbf{x}^* is a minimum. Hence by *reductio ad absurdum*, $v \geq 0$.

Define the $m' \times 1$ column vector $\underline{\mu}$ by $\underline{\mu}^T = (\underline{v}^T ; 0)$. Then rearrange $\underline{\mu}$ so that $\mu_i = 0$ if c_i is not active and $\mu_i \geq 0$ if c_i is active. As in sections 5.1 and 5.2 we can abbreviate this to

$$\underline{\mu}^T \underline{c}(\mathbf{x}^*) = 0 \quad (5-4-4)$$

because $\underline{\mu} \geq 0$ and $\underline{c}(\mathbf{x}^*) \leq 0$. Also

$$\frac{dc_A}{d\mathbf{x}} \underline{v} = \frac{dc_A}{d\mathbf{x}} (\mathbf{x}^*) \underline{v} + \frac{dc_B}{d\mathbf{x}} (\mathbf{x}^*) \underline{0} = \frac{dc}{d\mathbf{x}} (\mathbf{x}^*) \underline{\mu} . \quad (5-4-5)$$

Substituting (5-4-5) into (5-4-3) we have

$$\frac{df}{d\mathbf{x}} (\mathbf{x}^*) + \frac{dq}{d\mathbf{x}} (\mathbf{x}^*) \underline{\lambda} + \frac{dc}{d\mathbf{x}} (\mathbf{x}^*) \underline{\mu} = 0 .$$

We have thus derived the Kuhn-Tucker conditions. We see that the middle condition (5-4-4) is just an abbreviation of the restriction that the multipliers corresponding to non-active constraints must be zero. (5-4-4) is sometimes called the *complementary slackness condition*.

5.5 Second order conditions

As usual we state the conditions first. The second order condition that \underline{x}^* be a solution of G is that

$$\Delta \underline{x}^T \frac{d^2 f}{d \underline{x}^2} (\underline{x}^*) \Delta \underline{x} \geq 0$$

for all vectors $\Delta \underline{x}$ satisfying

$$\Delta \underline{x}^T \frac{d \underline{a}}{d \underline{x}} (\underline{x}^*) = \underline{0}^T,$$

where $\underline{a}(\underline{x})$ represents the vector of equation and active constraints.

The sufficient conditions are

$$\frac{d f}{d \underline{x}} (\underline{x}^*) = \underline{0}$$

and there exist $\underline{\mu} \geq \underline{0}$ such that

$$\underline{\mu}^T \underline{c}(\underline{x}^*) = 0$$

and

$$\Delta \underline{x}^T \frac{d^2 f}{d \underline{x}^2} (\underline{x}^*) \Delta \underline{x} > 0$$

for all sufficiently small vectors $\Delta \underline{x}$ satisfying

$$\Delta \underline{x}^T \frac{d \underline{a}}{d \underline{x}} (\underline{x}^*) = \underline{0}^T.$$

(As before, the sufficient conditions imply the existence of Lagrange multipliers.)

The proofs follow exactly those of the analogous conditions of the problem E , except that everywhere active inequality constraints are treated as equation constraints. The non-active constraints only occur in the Lagrangian, where they are multiplied by the zero entries of $\underline{\mu}$.

6 THE DUALITY THEOREM

We complete these notes with a statement and derivation of the *duality* (or *Kuhn-Tucker*) *theorem*. This important theorem underlies most numerical methods of constrained optimization.

6.1 The dual function

We define the *dual function* $\phi(\underline{\lambda}, \underline{\mu})$ for the general constrained problem G by

$$\phi(\underline{\lambda}, \underline{\mu}) = \min_{\underline{x}} \left\{ f(\underline{x}) + \underline{\lambda}^T \underline{g}(\underline{x}) + \underline{\mu}^T \underline{c}(\underline{x}) \right\}.$$

Let \underline{x}^* be the vector which minimizes

$$f(\underline{x}) + \underline{\lambda}^{*T} \underline{q}(\underline{x}) + \underline{\mu}^{*T} \underline{c}(\underline{x})$$

where $\underline{\mu}^* \geq \underline{0}$.

Note that

$$\phi(\underline{\lambda}^*, \underline{\mu}^*) = f(\underline{x}^*) + \underline{\lambda}^{*T} \underline{q}(\underline{x}^*) + \underline{\mu}^{*T} \underline{c}(\underline{x}^*).$$

6.2 Statement of the theorem

\underline{x}^* is a solution of G

(i) if $\phi(\underline{\lambda}^*, \underline{\mu}^*) \geq \phi(\underline{\lambda}, \underline{\mu})$ for all $\underline{\lambda}$ and for all $\underline{\mu} \geq \underline{0}$,

(ii) only if (provided the matrix $\partial^2 \mathcal{L} / \partial \underline{x}^2$ is everywhere positive definite) there exist $\underline{\lambda}^*, \underline{\mu}^*$ which maximize $\phi(\underline{\lambda}, \underline{\mu})$ for all $\underline{\lambda}$ and for all $\underline{\mu} \geq \underline{0}$.

6.3 Proof of part (i)

Let $\underline{\mu}^* \geq \underline{0}$ and $\underline{\lambda}^*$ be vectors which maximize $\phi(\underline{\lambda}, \underline{\mu})$ for all $\underline{\lambda}$ and for all $\underline{\mu} \geq \underline{0}$. Then $\underline{\lambda}^*, \underline{\mu}^*$ must satisfy the Kuhn-Tucker conditions for the maximization problem

$$\begin{array}{l} \text{maximize } \phi(\underline{\lambda}, \underline{\mu}) \text{ subject to } \underline{\mu} \geq \underline{0} \\ \underline{\lambda}, \underline{\mu} \end{array}$$

Denote the inequality constraints $\underline{\mu} \geq \underline{0}$ by $\underline{\gamma}(\underline{\mu}) = -\underline{\mu} \leq \underline{0}$. Let the multipliers associated with $\underline{\gamma}(\underline{\mu})$ be $\underline{\alpha}$. Then since there are no equation constraints, (5-4-1) becomes

$$\left. \begin{aligned} \frac{\partial \phi}{\partial \underline{\lambda}} + \frac{\partial \underline{\gamma}}{\partial \underline{\lambda}} \underline{\alpha} &= \underline{0} \\ \frac{\partial \phi}{\partial \underline{\mu}} + \frac{\partial \underline{\gamma}}{\partial \underline{\mu}} \underline{\alpha} &= \underline{0} \\ \underline{\alpha}^T \underline{\gamma}(\underline{\mu}) &= 0 \\ \underline{\alpha} &\geq \underline{0} \end{aligned} \right\} \quad (6-3-1)$$

and

Since $\frac{\partial \underline{\gamma}}{\partial \underline{\lambda}}(\underline{\mu}) = \underline{0}$ and $\frac{\partial \underline{\gamma}}{\partial \underline{\mu}} = \underline{1}$, and $\underline{\alpha} \geq \underline{0}$, (6-3-1) becomes

$$\left. \begin{aligned} \frac{\partial \phi}{\partial \lambda} &= \underline{0} \\ \frac{\partial \phi}{\partial \mu} &= -\underline{\alpha} \leq \underline{0} \\ \left(\frac{\partial \phi}{\partial \mu}\right)^T \mu &= 0 \end{aligned} \right\} \quad (6-3-2)$$

Let $\hat{\underline{x}}(\lambda, \mu)$ be the vector which minimizes $f(\underline{x}) + \lambda^T \underline{q}(\underline{x}) + \mu^T \underline{c}(\underline{x})$.

Then

$$\phi(\lambda, \mu) = f(\hat{\underline{x}}(\lambda, \mu)) + \lambda^T \underline{q}(\hat{\underline{x}}(\lambda, \mu)) + \mu^T \underline{c}(\hat{\underline{x}}(\lambda, \mu)).$$

Now

$$\begin{aligned} \frac{\partial \phi}{\partial \lambda} &= \frac{\partial f}{\partial \lambda} + \frac{\partial \lambda}{\partial \lambda} \underline{q} + \frac{\partial \underline{q}}{\partial \lambda} \lambda + \frac{\partial \mu}{\partial \lambda} \underline{c} + \frac{\partial \underline{c}}{\partial \lambda} \mu \\ &= \frac{\partial f}{\partial \lambda} + \underline{q} + \frac{\partial \underline{q}}{\partial \lambda} \lambda + \frac{\partial \underline{c}}{\partial \lambda} \mu \\ &= \frac{\partial \underline{x}}{\partial \lambda} \frac{df}{d\underline{x}} + \frac{\partial \underline{x}}{\partial \lambda} \frac{d\underline{q}}{d\underline{x}} \lambda + \frac{\partial \underline{x}}{\partial \lambda} \frac{d\underline{c}}{d\underline{x}} \mu + \underline{q} \end{aligned}$$

where we have used (A-2). Therefore

$$\frac{\partial \phi}{\partial \lambda} = \frac{d\underline{x}}{d\lambda} \left(\frac{df}{d\underline{x}} + \frac{d\underline{q}}{d\underline{x}} \lambda + \frac{d\underline{c}}{d\underline{x}} \mu \right) + \underline{q}(\hat{\underline{x}}).$$

At any point $\hat{\underline{x}}$ which minimizes \mathcal{L} we have $\frac{df}{d\underline{x}} + \frac{d\underline{q}}{d\underline{x}} \lambda + \frac{d\underline{c}}{d\underline{x}} \mu = \underline{0}$.

Therefore

$$\frac{\partial \phi}{\partial \lambda} = \underline{q}(\hat{\underline{x}}) \quad (6-3-3)$$

Similarly,

$$\frac{\partial \phi}{\partial \mu} = \underline{c}(\hat{\underline{x}}) \quad (6-3-4)$$

In particular, (6-3-3) and (6-3-4) must hold for $\underline{x}^* = \hat{\underline{x}}(\lambda^*, \mu^*)$.

Substituting them into (6-3-2) we obtain

$$\left. \begin{aligned} \underline{q}(\underline{x}^*) &= \underline{0} \\ \underline{c}(\underline{x}^*) &\leq \underline{0} \end{aligned} \right\} \quad (6-3-5)$$

$$\mu^T \underline{c}(\underline{x}^*) = 0. \quad (6-3-6)$$

(6-3-5) show that \underline{x}^* satisfies the constraints of the problem G.

Let \underline{x} be any point that also satisfies the constraints. Then by definition of $\phi(\underline{\lambda}^*, \underline{\mu}^*)$

$$f(\underline{x}^*) + \underline{\lambda}^{*T} \underline{q}(\underline{x}^*) + \underline{\mu}^{*T} \underline{c}(\underline{x}^*) \leq f(\underline{x}) + \underline{\lambda}^{*T} \underline{q}(\underline{x}) + \underline{\mu}^{*T} \underline{c}(\underline{x}) . \quad (6-3-7)$$

But $\underline{q}(\underline{x}^*)$ and $\underline{q}(\underline{x})$ are zero. From (6-3-6), $\underline{\mu}^{*T} \underline{c}(\underline{x}^*)$ is also zero. Since $\underline{c}(\underline{x}) \leq \underline{0}$ and $\underline{\mu}^* \geq \underline{0}$ we have $\underline{\mu}^{*T} \underline{c}(\underline{x}) \leq 0$. Hence we get

$$\begin{aligned} f(\underline{x}^*) &= f(\underline{x}^*) + \underline{\lambda}^{*T} \underline{q}(\underline{x}^*) + \underline{\mu}^{*T} \underline{c}(\underline{x}^*) \\ &\leq f(\underline{x}) + \underline{\lambda}^{*T} \underline{q}(\underline{x}) + \underline{\mu}^{*T} \underline{c}(\underline{x}) = f(\underline{x}) + \underline{\mu}^{*T} \underline{c}(\underline{x}) \leq f(\underline{x}) . \end{aligned}$$

Hence \underline{x}^* is a solution of the problem G.

6.4 Proof of part (ii)

Let \underline{x}^* be a solution of G and let $\partial^2 \mathcal{L} / \partial \underline{x}^2$ be positive definite everywhere.

At any point $\hat{\underline{x}}(\underline{\lambda}, \underline{\mu})$ which minimizes \mathcal{L} we must have

$$\frac{df}{d\underline{x}} + \frac{dq}{d\underline{x}} \underline{\lambda} + \frac{dc}{d\underline{x}} \underline{\mu} = \underline{0}$$

which can be written

$$\frac{df}{d\underline{x}} + \sum_{i=1}^m \lambda_i \frac{dq_i}{d\underline{x}} + \sum_{i=1}^{m'} \mu_i \frac{dc_i}{d\underline{x}} = \underline{0} . \quad (6-4-1)$$

Differentiating with respect to λ_j gives

$$\frac{\partial}{\partial \lambda_j} \left(\frac{df}{d\underline{x}} \right) + \sum_{i \neq j} \lambda_i \frac{\partial}{\partial \lambda_j} \left(\frac{dq_i}{d\underline{x}} \right) + \frac{\partial}{\partial \lambda_j} \left(\lambda_j \frac{dq_j}{d\underline{x}} \right) + \sum_{i=1}^m \mu_i \frac{\partial}{\partial \lambda_j} \left(\frac{dc_i}{d\underline{x}} \right) = \underline{0} . \quad (6-4-2)$$

In view of (A-3), (6-4-2) becomes

$$\left(\frac{d^2 f}{d\underline{x}^2} \right)^T \frac{\partial \underline{x}}{\partial \lambda_j} + \sum_{i \neq j} \lambda_i \left(\frac{d^2 q_i}{d\underline{x}^2} \right)^T \frac{\partial \underline{x}}{\partial \lambda_j} + \frac{dq_j}{d\underline{x}} + \lambda_j \left(\frac{d^2 q_j}{d\underline{x}^2} \right)^T \frac{\partial \underline{x}}{\partial \lambda_j} + \sum_{i=1}^m \mu_i \left(\frac{d^2 c_i}{d\underline{x}^2} \right)^T \frac{\partial \underline{x}}{\partial \lambda_j} = \underline{0}$$

which can be further simplified to

$$\left(\frac{d^2 \mathcal{L}}{d\underline{x}^2} \right)^T \frac{\partial \underline{x}}{\partial \lambda_j} + \frac{dq_j}{d\underline{x}} = \underline{0} . \quad (6-4-3)$$

For $j = 1, \dots, m$, the row vector equations (6-4-3) can be combined to obtain the matrix equation

$$\left(\frac{d^2 \mathcal{L}}{d\mathbf{x}^2} \right)^T \left(\frac{\partial \mathbf{x}}{\partial \lambda} \right)^T + \frac{d\mathbf{q}}{d\mathbf{x}} = 0 \quad (6-4-4)$$

Since $d^2 \mathcal{L}/d\mathbf{x}^2$ is positive definite it has an inverse. Hence (6-4-4) may be written

$$\frac{\partial \mathbf{x}}{\partial \lambda} = - \left(\frac{d\mathbf{q}}{d\mathbf{x}} \right)^T \left(\frac{d^2 \mathcal{L}}{d\mathbf{x}^2} \right)^{-1} \quad (6-4-5)$$

Similarly, by differentiating (6-4-1) by μ_j we obtain

$$\frac{\partial \mathbf{x}}{\partial \mu} = - \left(\frac{d\mathbf{c}}{d\mathbf{x}} \right)^T \left(\frac{d^2 \mathcal{L}}{d\mathbf{x}^2} \right)^{-1} \quad (6-4-6)$$

Putting

$$\underline{v} = \begin{pmatrix} \lambda \\ \dots \\ \mu \end{pmatrix} \quad \text{and} \quad \underline{\alpha} = \begin{pmatrix} q \\ \dots \\ c_A \end{pmatrix}$$

where c_A is the vector of active constraints, we have

$$\frac{d^2 \phi}{d\underline{v}^2} = \begin{pmatrix} \frac{\partial^2 \phi}{\partial \lambda^2} & \dots & \frac{\partial^2 \phi}{\partial \lambda \partial \mu} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \phi}{\partial \mu \partial \lambda} & \dots & \frac{\partial^2 \phi}{\partial \mu^2} \end{pmatrix}$$

But from (6-3-3) and (6-3-4) we have

$$\frac{\partial^2 \phi}{\partial \lambda^2} = \frac{\partial}{\partial \lambda} \left(\frac{\partial \phi}{\partial \lambda} \right) = \frac{\partial}{\partial \lambda} \left[\mathbf{g}(\hat{\mathbf{x}}(\lambda, \mu)) \right] = \frac{\partial \mathbf{x}}{\partial \lambda} \frac{d\mathbf{q}}{d\mathbf{x}} \quad (6-4-7)$$

and

$$\frac{\partial^2 \phi}{\partial \lambda \partial \mu} = \frac{\partial}{\partial \lambda} \left(\frac{\partial \phi}{\partial \mu} \right) = \frac{\partial}{\partial \lambda} \left[\mathbf{c}(\hat{\mathbf{x}}(\lambda, \mu)) \right] = \frac{\partial \mathbf{x}}{\partial \lambda} \frac{d\mathbf{c}}{d\mathbf{x}} \quad (6-4-8)$$

Similarly we have

$$\frac{\partial^2 \phi}{\partial \mu \partial \lambda} = \frac{\partial \mathbf{x}}{\partial \mu} \frac{d\mathbf{q}}{d\mathbf{x}} \quad (6-4-9)$$

and

$$\frac{\partial^2 \phi}{\partial \mu^2} = \frac{\partial \mathbf{x}}{\partial \mu} \frac{d\mathbf{c}}{d\mathbf{x}} \quad (6-4-10)$$

Therefore

$$\frac{d^2\phi}{d\underline{v}^2} = \begin{pmatrix} \frac{\partial \underline{x}}{\partial \underline{\lambda}} \frac{d\underline{q}}{d\underline{x}} : \frac{\partial \underline{x}}{\partial \underline{\lambda}} \frac{d\underline{c}}{d\underline{x}} \\ \dots\dots\dots : \dots\dots\dots \\ \frac{\partial \underline{x}}{\partial \underline{\mu}} \frac{d\underline{q}}{d\underline{x}} : \frac{\partial \underline{x}}{\partial \underline{\mu}} \frac{d\underline{c}}{d\underline{x}} \end{pmatrix}.$$

Substituting in (6-4-5) and (6-4-6) we get

$$\frac{d^2\phi}{d\underline{v}^2} = \begin{pmatrix} -\left(\frac{d\underline{q}}{d\underline{x}}\right)^T \left(\frac{d^2f}{d\underline{x}^2}\right)^{-1} \frac{d\underline{q}}{d\underline{x}} : -\left(\frac{d\underline{c}}{d\underline{x}}\right)^T \left(\frac{d^2f}{d\underline{x}^2}\right)^{-1} \frac{d\underline{c}}{d\underline{x}} \\ \dots\dots\dots : \dots\dots\dots \\ -\left(\frac{d\underline{q}}{d\underline{x}}\right)^T \left(\frac{d^2f}{d\underline{x}^2}\right)^{-1} \frac{d\underline{q}}{d\underline{x}} : -\left(\frac{d\underline{c}}{d\underline{x}}\right)^T \left(\frac{d^2f}{d\underline{x}^2}\right)^{-1} \frac{d\underline{c}}{d\underline{x}} \end{pmatrix}.$$

Hence

$$\frac{d^2\phi}{d\underline{v}^2} = -\left(\frac{d\underline{a}}{d\underline{x}}\right)^T \frac{d^2f}{d\underline{x}^2} \frac{d\underline{a}}{d\underline{x}}. \quad (6-4-11)$$

Since $d\underline{a}/d\underline{x}$ is of full rank and $d^2f/d\underline{x}^2$ is positive definite, (6-4-11) implies that $d^2\phi/d\underline{v}^2$ is negative definite.

Define

$$\underline{v}^* = \begin{pmatrix} \underline{\lambda}^* \\ \dots\dots\dots \\ \underline{\mu}^* \end{pmatrix}$$

to be the Lagrange multipliers corresponding to \underline{x}^* . We shall show that \underline{v}^* satisfies all the sufficient conditions to maximize $\phi(\underline{v})$ for all $\underline{\mu} \geq \underline{0}$ and for all $\underline{\lambda}$.

Let $\underline{\gamma}$ be the Lagrange multipliers of this problem. Then the Lagrangian is $L = \phi - \underline{\gamma}^T \underline{\mu}$. Now from (6-3-3) and (6-3-4) we have that

$$\frac{\partial L}{\partial \underline{\lambda}} = \frac{\partial \phi}{\partial \underline{\lambda}} = \underline{q}(\underline{x}^*) = \underline{0}$$

$$\frac{dL}{d\underline{\mu}} = \frac{\partial \phi}{\partial \underline{\mu}} - \underline{\gamma} = \underline{c}(\underline{x}^*) - \underline{\gamma}. \quad (6-4-12)$$

We set $\gamma_i = 0$ if $c_i(\underline{x}^*) = 0$ and $\gamma_i = -c_i(\underline{x}^*)$ if $c_i(\underline{x}^*) < 0$. Then $\underline{\gamma} \geq \underline{0}$. Also, because $\underline{\mu}$ is itself a Kuhn-Tucker multiplier, we have that $\mu_i = 0$ if $c_i(\underline{x}^*) = 0$ and $\mu_i > 0$ if $c_i(\underline{x}^*) < 0$. Hence $\gamma_i = 0$ if $\mu_i = 0$ and $\gamma_i > 0$ if $\mu_i > 0$. We can express this as $\underline{\gamma}^T \underline{\mu} = 0$. From (6-4-12) we therefore have that $dL/d\underline{\mu} = \underline{c}(\underline{x}^*) - \underline{\gamma} = \underline{0}$. Hence $dL/d\underline{\mu} = \underline{0}$ and therefore \underline{v}^* satisfies the sufficient conditions and we have finished our proof.

Appendix A

STANDARD RESULTS OF VECTOR DIFFERENTIATION

Apart from the results discussed in section 2.2, the following are also used in these notes.

(1) Let $\underline{x}^T = \begin{pmatrix} \underline{v}^T & \underline{u}^T \end{pmatrix}$ and suppose $f = f(\underline{u}, \underline{v})$.

Then the second *total* derivative of f along the contour $\underline{v} = \underline{v}(\underline{u})$ is given by

$$\frac{d^2 f}{d\underline{u}^2} = \frac{d\underline{x}}{d\underline{u}} \frac{d^2 f}{d\underline{x}^2} \left(\frac{d\underline{x}}{d\underline{u}} \right)^T + \sum_j \frac{\partial f}{\partial v_j} \frac{d^2 v_j}{d\underline{u}^2}. \quad (\text{A-1})$$

(2) Let $f = f(\underline{x})$ and suppose $\underline{x} = \underline{x}(\underline{u})$, for some vector \underline{u} . A *chain rule* applies in the form

$$\frac{df}{d\underline{u}} = \frac{d\underline{x}}{d\underline{u}} \frac{df}{d\underline{x}}. \quad (\text{A-2})$$

(3) Let \underline{f} be an m -vector function of n variables \underline{x} . Let the \underline{x} also depend on a scalar t . Then \underline{f} is implicitly a function of t and the chain rule is

$$\frac{d\underline{f}}{dt} = \left(\frac{d\underline{f}}{d\underline{x}} \right)^T \frac{d\underline{x}}{dt}. \quad (\text{A-3})$$

PRECEDING PAGE BLANK-NOT FILMED

Appendix B

STATEMENTS OF THEOREMS ASSUMED IN THE TEXT

B.1 Implicit function theorem³

Let $\underline{u}(\underline{x})$ be m continuously differentiable functions of n variables \underline{x} ($m < n$).
If

$$\text{rank} \left(\frac{d\underline{u}}{d\underline{x}} \right) = m$$

then it is possible to solve for m of the variables, say x_1, \dots, x_m , in terms of the remaining $n - m$ variables x_{m+1}, \dots, x_n ,
i.e.

$$x_i = \psi_i(x_{m+1}, \dots, x_n) \quad (i = 1, \dots, m).$$

The m functions $\underline{\psi}$ are called *implicit functions*.

B.2 Mean value theorem

Both the mean value theorem and the second mean value theorem can be derived from the *fundamental inequality*³.

Let f be a differentiable function of n variables \underline{x} . Then there exists a number ξ satisfying $0 < \xi < 1$ such that

$$\Delta \underline{x}^T \frac{df}{d\underline{x}} (\underline{x} + \xi \Delta \underline{x}) = f(\underline{x} + \Delta \underline{x}) - f(\underline{x}).$$

B.3 Second mean value theorem

Let f be a twice differentiable function of n variables \underline{x} . Then for all $\xi > 0$

$$\frac{df}{d\underline{x}} (\underline{x} + \xi \Delta \underline{x}) = \frac{df}{d\underline{x}} (\underline{x}) + \xi \frac{d^2 f}{d\underline{x}^2} (\underline{x}) \Delta \underline{x} + O(\xi^2).$$

B.4 Taylor's theorem for vector functions of vectors

Let \underline{f} be a column of m differentiable functions of an n -vector \underline{x} .
Then

$$\underline{f}(\underline{x} + \Delta \underline{x}) = \underline{f}(\underline{x}) + \left(\frac{d\underline{f}}{d\underline{x}} \right)^T \Delta \underline{x} + O|\Delta \underline{x}|^2.$$

B.5 Taylor's theorem for scalar functions of vectors³

Let f be a scalar valued and at least twice differentiable function of an n -vector \underline{x} . Then

$$f(\underline{x} + \Delta \underline{x}) = f(\underline{x}) + \left(\frac{df}{d\underline{x}} \right)^T \Delta \underline{x} + \frac{1}{2} \Delta \underline{x}^T \frac{d^2 f}{d\underline{x}^2} \Delta \underline{x} + O|\Delta \underline{x}|^3.$$

Note that

$$\left(\frac{df}{d\underline{x}}\right)^T \Delta \underline{x} = \Delta \underline{x}^T \frac{df}{d\underline{x}} .$$

B.6 The intermediate value theorem⁵

We state the n-dimensional analogue of this well-known theorem⁵ in the following form.

Let f be a continuous function of n variables \underline{x} . Let $f(\underline{x}_1) < 0$ and $f(\underline{x}_2) > 0$. Then there exists a point $\underline{\xi}$ lying on the line segment joining \underline{x}_1 and \underline{x}_2 such that $f(\underline{\xi}) = 0$.

LIST OF SYMBOLS

$\underline{a}(\underline{x})$	vector of active constraints
(a_{ij})	matrix whose i,j -entry is a_{ij}
$(A)_{ij}$	the i,j -entry of matrix A
$\underline{c}(\underline{x})$	vector of m' inequality constraints
$\frac{df}{d\underline{x}}$	the gradient vector, <i>eg</i> the column vector of first partial derivatives of the scalar function f
$\frac{d^2f}{d\underline{x}^2}$	the Hessian matrix, <i>eg</i> the matrix of second partial derivatives of the scalar function f
$f(\underline{x})$	a scalar function of the n variables \underline{x}
$\underline{f}(\underline{x})$	a vector-valued function of the n variables \underline{x}
I_n	the $n \times n$ identity matrix
L	the Lagrangian of the dual function
\mathcal{L}	the Lagrangian function $\mathcal{L} = f + \underline{\lambda}^T \underline{q}$ for the equation constrained problem $\mathcal{L} = f + \underline{\lambda}^T \underline{q} + \underline{\mu}^T \underline{c}$ for the general constrained problem
$\underline{0}$	column vector of zero entries
$\underline{0}^T$	row vector of zero entries
$O(\Delta x^n)$	terms of order Δx^n and higher terms
$O \Delta \underline{x} ^n$	terms of order $ \Delta \underline{x} ^n$ and higher terms
$[0,1]$	the closed interval of numbers between 0 and 1, <i>ie</i> the interval of numbers x such that $0 \leq x \leq 1$
$\underline{q}(\underline{x})$	vector of m equation constraints
T	symbol of vector or matrix transposition
\underline{x}	column vector of (not necessarily n) variables x_1, \dots, x_n
\underline{x}^*	solution of the particular minimization problem under discussion
$\underline{\lambda}$	column vector of Lagrange multipliers
$\underline{\lambda}^*$	column vector of Lagrange multipliers corresponding to \underline{x}^* in duality theorem
$\underline{\mu}$	column vector of Kuhn-Tucker multipliers
$\underline{\mu}^*$	column vector of Kuhn-Tucker multipliers corresponding to \underline{x}^* in duality theorem
Φ	the dual function $= \min_{\underline{x}} \left\{ f(\underline{x}) + \underline{\lambda}^T \underline{q}(\underline{x}) + \underline{\mu}^T \underline{c}(\underline{x}) \right\}$
\triangleq	definition symbol, <i>ie</i> the left hand side of the equation is <i>defined</i> by the right hand side

REFERENCES

- | <u>No.</u> | <u>Author</u> | <u>Title, etc</u> |
|------------|-------------------------------|---|
| 1 | D.G. Luenberger | Introduction to linear and nonlinear programming.
Reading, Massachusetts, Addison-Wesley (1973) |
| 2 | L.C.W. Dixon | Nonlinear optimisation.
London, The English Universities Press (1972) |
| 3 | M.D. Intrilligator | Mathematical optimization and economic theory.
Englewood Cliffs, New Jersey, Prentice-Hall Inc. (1971) |
| 4 | E. Kreyszig | Advanced Engineering Mathematics.
John Wiley & Sons, Inc. (1972) |
| 5 | W. Rudin | Principles of mathematical analysis.
McGraw-Hill Inc. (1964) |
| 6 | A.V. Fiacco
G.P. McCormick | Nonlinear programming: sequential unconstrained minimization
techniques.
Wiley, New York (1968) |

Figs 1&2

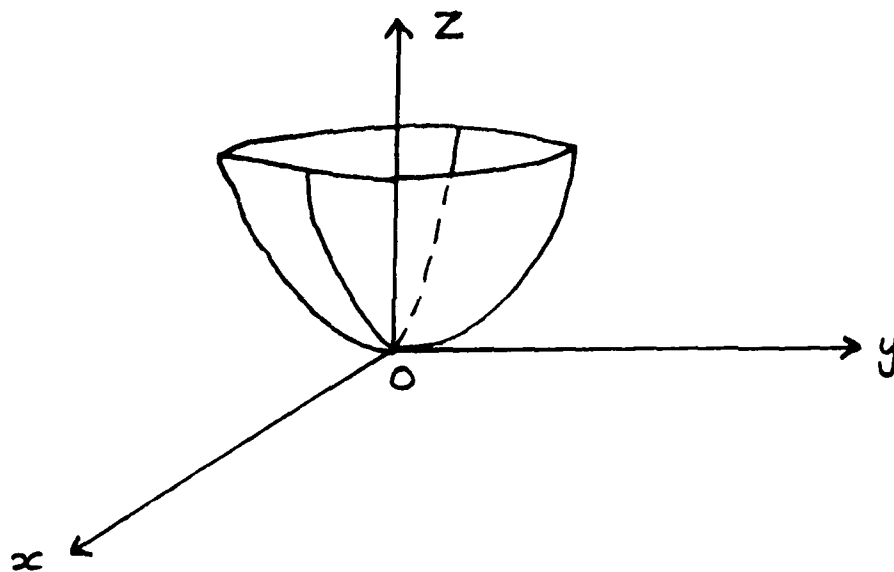


Fig 1 The paraboloid $z = x^2 + y^2$

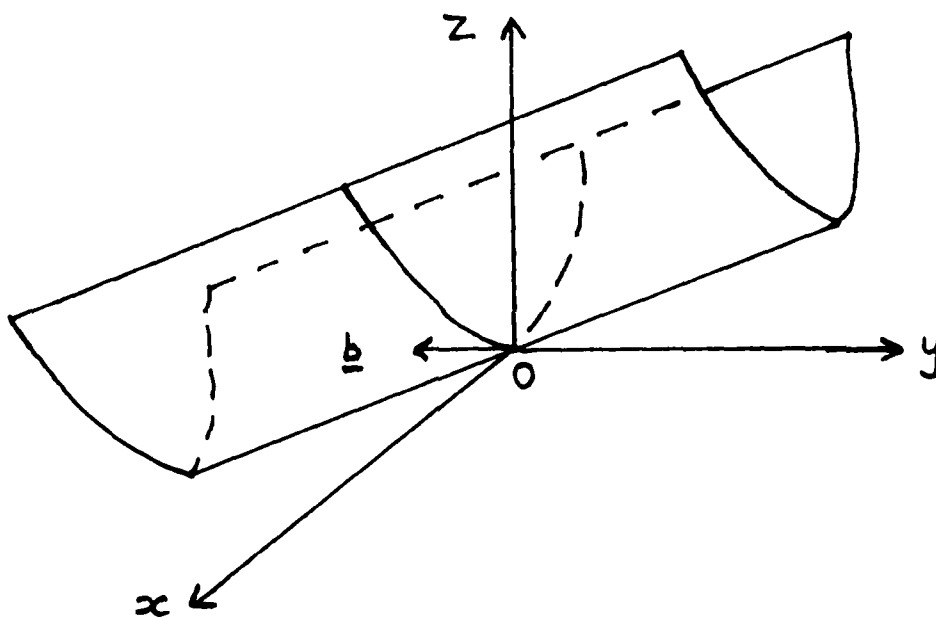


Fig 2 The surface $z = x^2 + y$

Figs 3&4

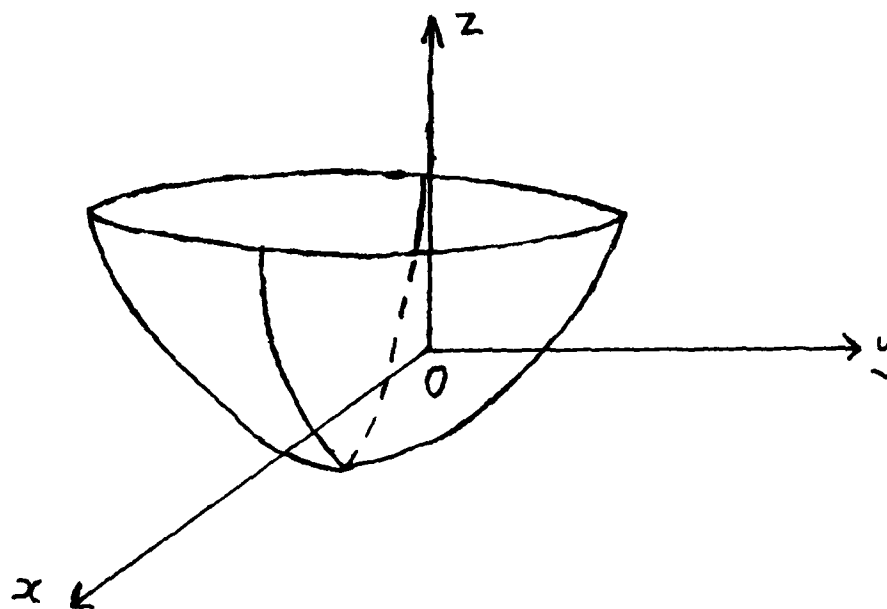


Fig 3 The paraboloid $z = x^2 + y^2 + y$

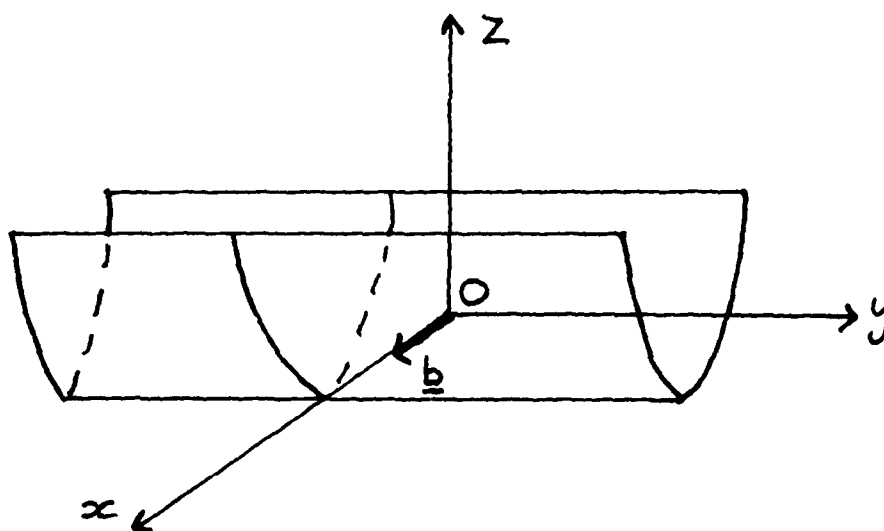


Fig 4 The surface $z = x^2 - x$

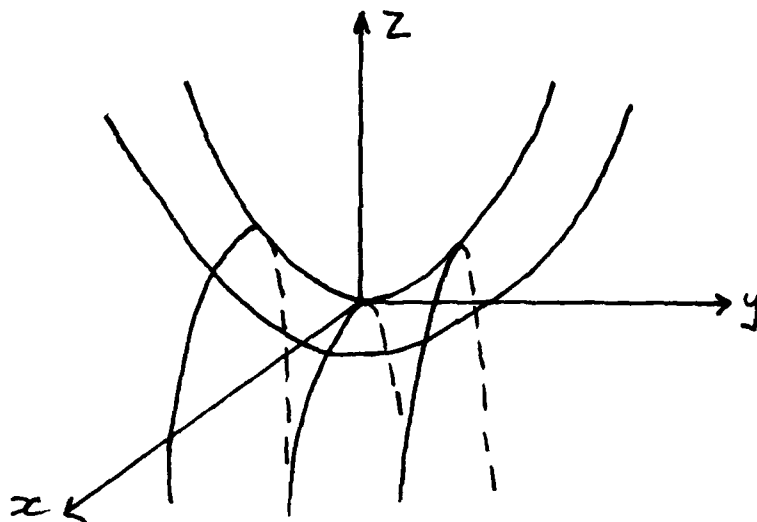


Fig 5 The saddle-shaped surface $z = y^2 - x^2$

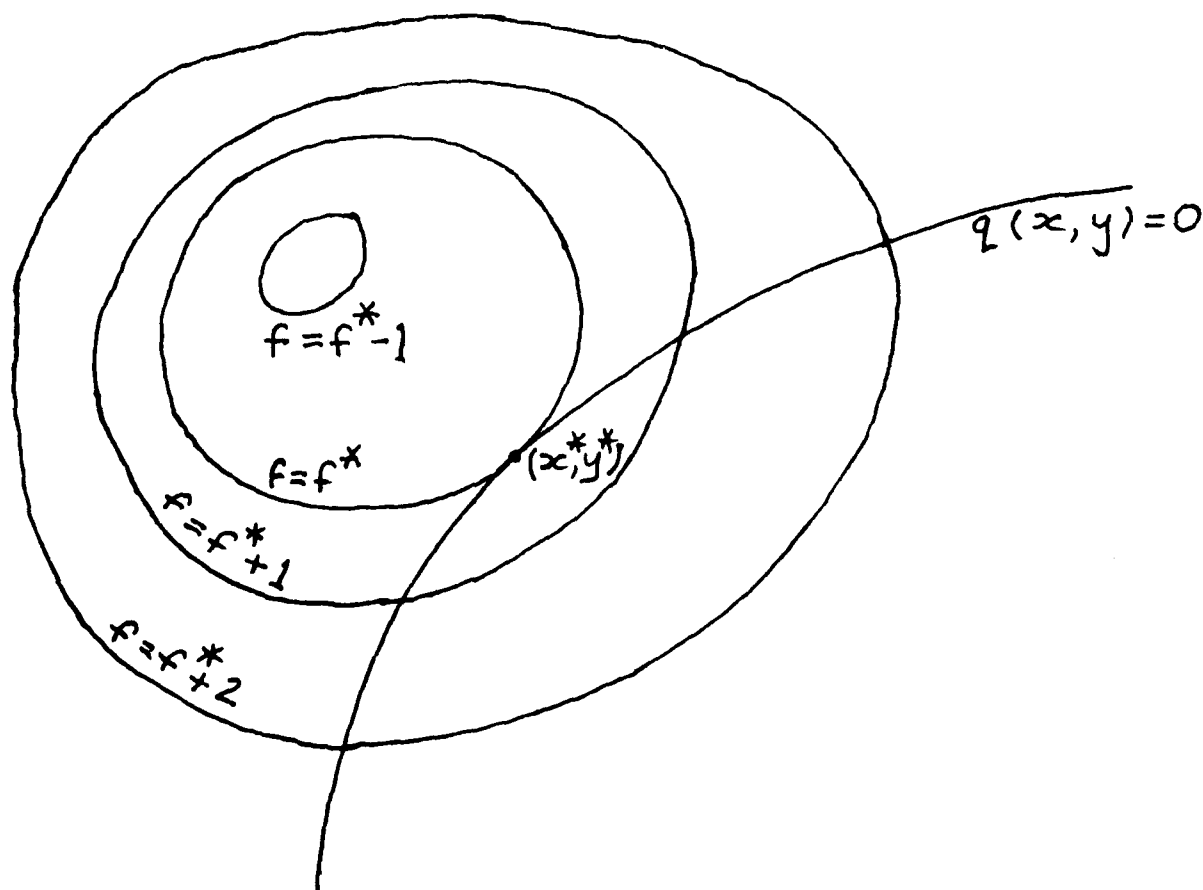


Fig 6 Contours of $f(x, y) = \text{const}$ in problem E2

Fig 7

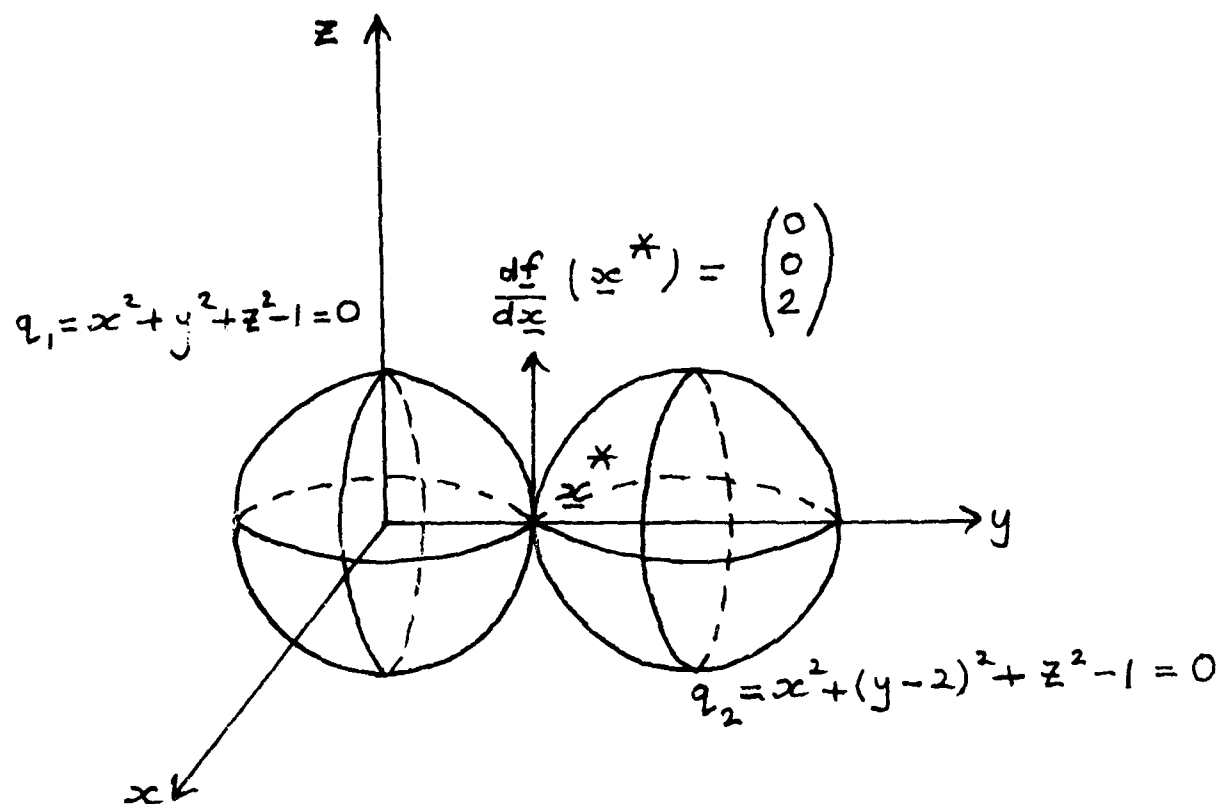


Fig 7 The gradient vector of $f = x^2 + (y-1)^2 + (z+1)^2$ at the intersection of the two spheres $q_1 = 0$ and $q_2 = 0$.

Fig 8

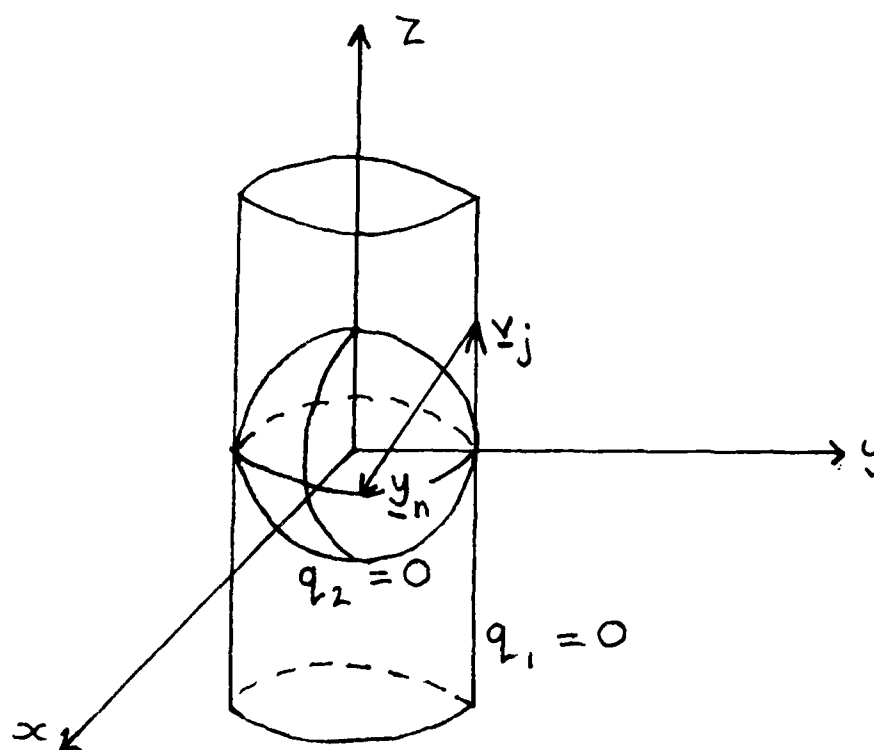


Fig 8 The constraints $q_1 = x^2 + y^2 - 1 = 0$ and $q_2 = x^2 + y^2 + z^2 - 1 = 0$, as well as a typical vector \underline{y}_n .

REPORT DOCUMENTATION PAGE

Overall security classification of this page

UNLIMITED

As far as possible this page should contain only unclassified information. If it is necessary to enter classified information, the box above must be marked to indicate the classification, e.g. Restricted, Confidential or Secret.

1. DRIC Reference (to be added by DRIC)	2. Originator's Reference RAE TM Aero 1888	3. Agency Reference N/A	4. Report Security Classification/Marking UNLIMITED		
5. DRIC Code for Originator 7673000W		6. Originator (Corporate Author) Name and Location Royal Aircraft Establishment, Farnborough, Hants, UK			
5a. Sponsoring Agency's Code N/A		6a. Sponsoring Agency (Contract Authority) Name and Location N/A			
7. Title Necessary and sufficient conditions in nonlinear optimization					
7a. (For Translations) Title in Foreign Language					
7b. (For Conference Papers) Title, Place and Date of Conference					
8. Author 1. Surname, Initials Skrobanski, J.J.	9a. Author 2	9b. Authors 3, 4		10. Date January 1981	Pages 51
				Refs. 6	
11. Contract Number N/A	12. Period N/A	13. Project		14. Other Reference Nos.	
15. Distribution statement (a) Controlled by – (b) Special limitations (if any) – Unlimited					
16. Descriptors (Keywords) (Descriptors marked * are selected from TEST) Optimization.					
17. Abstract These notes are directed at the newcomer to nonlinear programming for whom a thorough understanding of Lagrange multipliers, the Kuhn-Tucker conditions and the duality theorem is essential. The notes attempt to explain these foundations of the theory and what motivates them. Special cases of one or two dimensions are considered and are extended by means of the notation of vector differentiation to the case of n variables. The reader is taken in stages from the problem of unconstrained minimization, through the equation constrained problem, to the general constrained problem. The important Jacobian assumption is also discussed.					

**DATA
FILM**